Vibrating vaginal balls to improve pelvic floor muscle performance in women after childbirth: A randomised controlled feasibility trial

Claudia Oblasser

Thesis submitted for the degree of

Doctor of Philosophy (PhD) in Midwifery

City, University of London

School of Health Sciences

Centre for Maternal and Child Health Research

September 2018

Volume I

CONTENT

LIST OF TABLES

LIST OF BOXES

LIST OF FIGURES

DEDICATION

With gratitude, I dedicate this thesis to my parents who supported me to begin a path of higher education despite this not being common for girls in our Austrian village in the 1970s.

“Dubium sapientiae initium” (Doubt is the origin of wisdom) (René Descartes),

and

“If I have the gift of prophecy and can fathom all mysteries and all knowledge, and if I have a faith that can move mountains, but do not have love, I am nothing” (1 Cor 13: 2).

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my research participants, without whom this thesis would not exist. Thank you for your interest, time and effort, and for sharing your experiences.

Then, my thanks for support, advice and encouragement go to

- my supervision team: Professor Dr. Christine McCourt, Dr. Janice Christie and Dr. Shashivadan Hirani;
- my Austrian advisor, gatekeeper, formal principal investigator and outcome assessor, Professor Engelbert Hanzal MD; and to
- my outcome assessors Edona Berisha BSc MSc MSc RM and Sabine Clauss MD.

I am also hugely indebted to the following persons without whose support I would not have been able to do this work:

- my midwifery colleagues, in particular those who supported recruitment;
- my obstetrician colleagues who supported recruitment;
- my physiotherapy colleagues;
- my fellow PhD students at City, especially Miriam Ashford née Thiel, Susan Bradley, and Nathalie Leister;
- all those who facilitated my fieldwork;
- The Analysis Factor statistics company, the Section of Medical Statistics (IMS) of the Medical University of Vienna/Austria, and statistician Wolfgang Peter; and
- my many friends who were always available for questions on language, layout, mathematical, physical, technical and IT matters, and who supported me practically and emotionally.

My special thanks go to

- my husband Thomas for 21 years of love and encouragement, and to
- my friend Sharky for love and support in London.

This thesis would not have been possible without the Doctoral Scholarship I received from City University London. Extra funds for travel and conference costs were provided by a School of Health Sciences Conference & Travel Grant and two travel bursaries from Santander and the Worshipful Company of Saddlers.

DECLARATION

I, Claudia Oblasser, confirm that the work presented in this thesis is my own.

I hereby grant powers of discretion to the University Librarian to allow the thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

30 September 2018

ABSTRACT

Background
Vibrating vaginal balls are available in Austria and other countries to enhance women's pelvic floor muscles following childbirth. There is currently little research evidence to assess their relative effectiveness in comparison to current standard care, which involves pelvic floor muscle exercises.

Aim
To assess practical issues and feasibility for optimal design of a future randomised controlled trial (RCT) which shall determine the efficacy of vibrating vaginal balls to improve pelvic floor muscle performance in postpartum women; to monitor potential harms of the experimental intervention and explore women's perspectives on and experiences with the interventions and the trial.

Design
Single (assessor) blind, randomised controlled feasibility RCT with two parallel groups.

Methods
Women after vaginal birth in Vienna were randomised into one of two intervention groups to use either a vibrating vaginal ball or the comparator pelvic floor muscle exercises for 12 weeks. Primary outcomes were feasibility criteria and necessary resources, secondary outcomes preliminary effect/harm and women's views and experiences. Data were analysed by statistics and content analysis, for effect by modified intention-to-treat and per protocol with Welch's *t*-test, for harms descriptively per protocol.

Results
134 women were screened, 56 randomised (35/18 into experimental/comparison group), 134 and 53-56 respectively analysed for feasibility objectives. The recruitment rate was 48.3%, 95% CI [39.2, 57.4], the adherence rate at best 62.9%, 95% CI [46.9, 78.9]. Change score difference for participant reported pelvic floor muscle strength was -5.1%, 95% CI [-18.0, 7.8], for maximum perineometric squeeze strength pressure 4.6 cm H_2O, 95% CI [-0.3, 9.4]. For vaginal balls, the risk for potential harm in the form of vulvovaginal symptoms was 13.5%, 95% CI [4.5, 28.8], for local discomfort 33.3%, 95% CI [17.9, 48.7]. Participants' opinion on the trial was encouraging.

Conclusion
A full RCT seems feasible with modifications. Public and patient involvement needs to clarify women's opinion on such a trial in the light of the feasibility findings.

Keywords
feasibility studies, randomised controlled trial, pelvic floor, vaginal balls/cones, resistance training, pelvic floor muscle performance, urinary incontinence, postnatal care, postpartum period

ACRONYMS AND ABBREVIATIONS

Abbreviations used in this thesis are listed here unless the abbreviation is well known, used fewer than three times and in close proximity, or used only in figures/tables in which case the abbreviation is defined in the respective legend.

AKH Vienna	Allgemeines Krankenhaus der Stadt Wien (Vienna General Hospital)
BMI	Body mass index
CE	Conformité européenne (explained in Footnote 91)
CI	Confidence interval
CO	Claudia Oblasser
EB	Edona Berisha
EH	Engelbert Hanzal
ICIQ-UI SF	International Consultation on Incontinence Modular Questionnaire – Urinary Incontinence Short Form
ID	Participant identifier
ITT	Intention to treat
mITT	Modified intention to treat
MSID	Minimal scientifically important difference
N, *n*	Number of participants
n.a.	Not applicable
NHS	UK National Health Service
NICE	National Institute for Health and Care Excellence
OR	Odds ratio
PI	Principal investigator
PP	Per protocol
PPI	Patient and public involvement
PRO	Participant reported outcome
PROM	Participant reported outcome measure
RCT	Randomised controlled trial
RR	Relative risk
SC	Sabine Clauss
SD	Standard deviation
vs	versus

1 INTRODUCTION

To familiarise the reader with the present work, this introductory chapter first explains the research problem and relevance of the topic. It then clarifies its connection to the profession of midwifery, thereby emphasising midwifery in Austria. It points out the research interest, gives an introduction to the methodology used and provides an overview of the thesis. Finally, the research team and study setting are introduced.

1.1 Research problem and relevance of the topic

Childbearing challenges female pelvic floor[1] integrity. This challenge of the inferior muscular closure of the bony pelvis may or may not lead to clinical symptoms (Baessler and Schüssler, 2008). Symptoms indicative of a possible pelvic floor impairment are urinary or anal incontinence, pelvic organ prolapse, or sexual problems (Baessler et al., 2008, Salvatore et al., 2017). The prevalence of these symptoms after childbirth can be as high as 38% for urinary incontinence, 17% for anal incontinence, and 7% for pelvic organ prolapse (Mørkved and Bø, 1999, Lukacz et al., 2006). Women after childbirth can experience a range of sexual difficulties (Brown and Lumley, 1998, Barrett et al., 2000, Abdool et al., 2009b, Leeman and Rogers, 2012, McDonald et al., 2015) of which some may originate from changes of the pelvic floor (Handa et al., 2008, Salvatore et al., 2017). Pelvic problems impair women's quality of life, women feel in a vulnerable situation and experience powerlessness in living with an unpredictable body; this negatively affects their intimate relationships, work and social activities (Margalith et al., 2004, Hägglund and Ahlström, 2007, O'Reilly et al., 2009). Pelvic dysfunction furthermore poses a financial burden to the health system (Imamura et al., 2010).

Pelvic floor muscle training is an effective nonsurgical treatment method for urinary incontinence generally[2] and also used in the peripartum time (Dumoulin et al., 2014, Woodley et al., 2017). In the sense of secondary prevention[3], pelvic floor muscle training can also aim at enhancing performance of the pelvic floor muscles in childbearing women without symptoms. It is then used with the intention to prevent urinary incontinence or other pelvic floor impairment symptoms which have been shown to occur as a consequence of

[1] As the thesis is about pelvic floor health in childbearing women, the term pelvic floor in the following implicitly denotes the female pelvic floor.
[2] For women in general, Luginbuehl et al. (2015a) recommend pelvic floor muscle training also for faecal incontinence, and scientific evidence seems to suggest effectiveness in the treatment of pelvic organ prolapse (Dumoulin et al., 2016).
[3] Secondary prevention is the interruption of a disease process before signs or symptoms of the disease appear (Farlex, 2018m).

pregnancy and birth (Milsom et al., 2017). As such, it is routinely recommended in postpartum care (Royal College of Midwives, 2009, The Joanna Briggs Institute, 2011, Tacke and Stüwe, 2013, Harder et al., 2014). Dumoulin et al. (2017), considering trials in which training started after childbirth, temper this statement by recommending that healthcare providers should carefully consider the cost/benefit of population based approaches to health professional taught pelvic floor muscle training to all postpartum women regardless of their continence status. Only recently, pelvic floor muscle training has come to be suggested as a routine intervention for all women during pregnancy[4] (National Institute for Health and Clinical Excellence, 2008/2017, Dumoulin et al., 2017, Woodley et al., 2017).

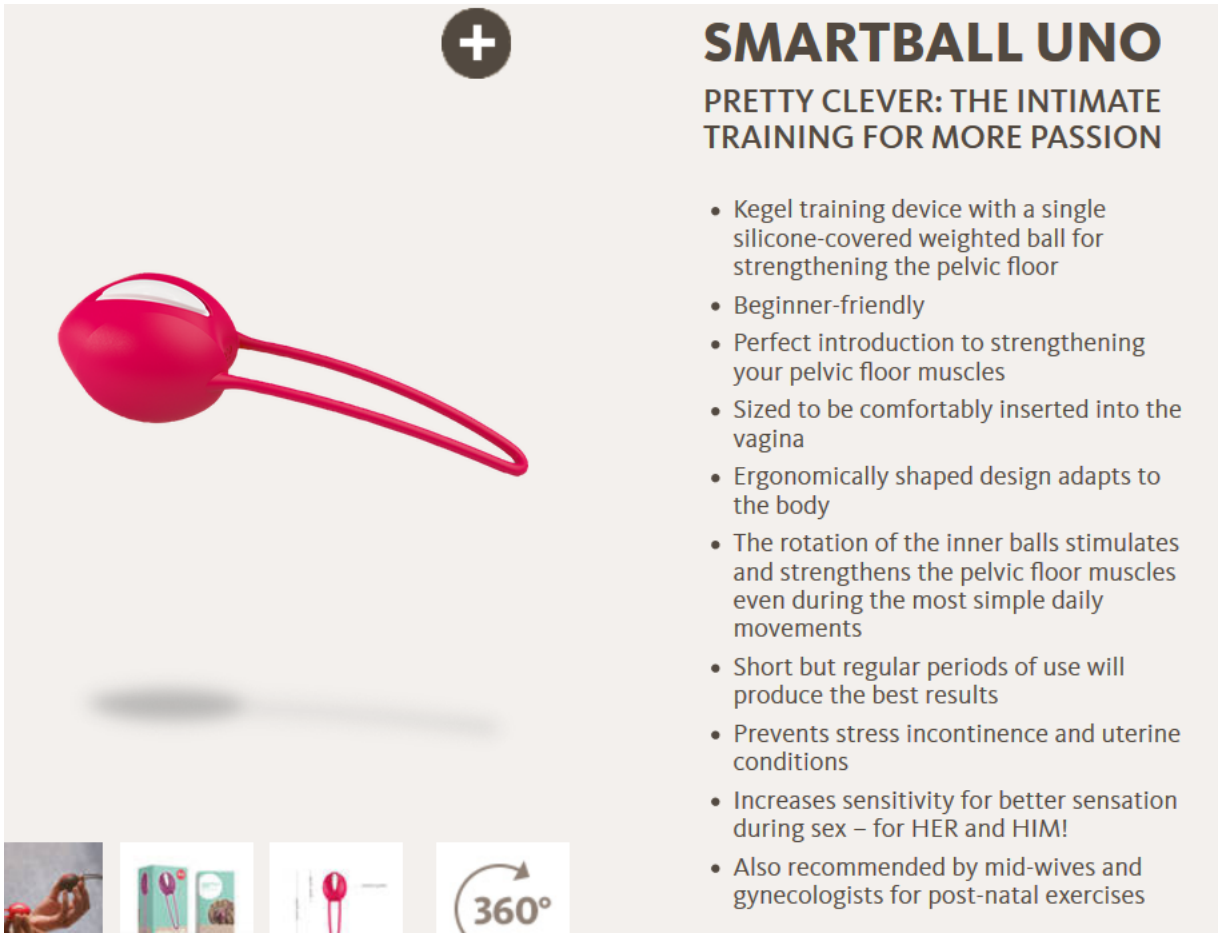
Pelvic floor muscle training is usually done by pelvic floor muscle exercises (Bø et al., 2015), but vibrating vaginal balls are available for over-the-counter sale in Austria and other countries as a device to strengthen the pelvic floor muscles after childbirth. These are balls with a smaller inner ball, and a popular method of use is to insert the ball vaginally and keep it in for 30 minutes once daily while moving around (FUN FACTORY, no date-d). It is said that the weight of the ball and the mechanical vibrations[5] would strengthen the pelvic floor muscles (Heller, 2015, ELANEE, 2017b, FUN FACTORY, no date-d, medesign I.C. GmbH, no date-b).[6] As a form of self-help treatment in a clinical self-management area, vibrating vaginal balls have been available for sale for a number of years. They were already mentioned by Buchheit (1985), and the FUN FACTORY company confirmed the production of such balls since approximately 1999 (FUN FACTORY, 2014).[7]

FUN FACTORY (no date-d, htpps://www.funfactory.com/en/smartballs/smartball-uno/) uses the slogan “Recommended by mid-wives and gynecologists” on its information web page (Figure 1). Similarly, the NUK pelvic floor trainer, a vibrating dumbbell shaped device (MAPA GmbH, no date-b), is advertised with reference to a survey among midwives and mothers. The company states that 85% of the 59 participating midwives think that this pelvic floor trainer is strengthening the pelvic floor muscles when used consequently, and that 88% of 27 mothers find its application comfortable and 81% would continue to use it. When attending a professional healthcare event around childbearing in Austria or Germany, one may find a stall selling vibrating vaginal balls for pelvic floor muscle rehabilitation. Likewise,

[4] As the thesis topic is postpartum pelvic floor muscle rehabilitation, information on pelvic floor care during pregnancy is kept to a minimum.
[5] Mechanical vibration as “continuing motion, often repetitive and periodic, of […] structures” ('Mechanical vibration', 2003, https://encyclopedia2.thefreedictionary.com/mechanical+vibration).
[6] With FUN FACTORY (no date-a) stating that the ball would even be of greater use than active contractions.
[7] For further historical information see Appendix A.

Austrian midwives and physiotherapists might inform women about the balls as an (auxiliary) preventive or therapeutic option. Vibrating vaginal balls are also mentioned in midwifery and physiotherapy teaching literature (Birk, 2012, Heller, 2015), and a physiotherapy student chose their use for preventive pelvic floor muscle strengthening as the topic of her final educational assignment (Butej, 2010).

Figure 1 Screenshot of the website of a well known vaginal ball manufacturer



Source: FUN FACTORY (no date-d). Item reprinted with permission of FUN FACTORY GmbH.

A device which helps to rehabilitate the pelvic floor muscles effectively without deliberate training would be practical and time saving as pelvic floor muscle exercises require motivation, adherence[8] and time (Mantle and Versi, 1991, Bø and Larsen, 1992, Ashworth and Hagan, 1993, Lagro-Janssen et al., 1994, Moore et al., 2013, Bø, 2015b). To women

[8] Adherence denotes the "extent to which the patient's behaviour matches agreed recommendations from the prescriber" (Horne et al., 2005, p. 12) and is the term recommended to be used for this concept in the context of a scientific agenda.

with a baby who are overtired and overbusy (Barclay et al., 1997), “passive exercise” without effort may sound attractive (Royal College of Midwives, 2018). The device might thus lead to a higher intervention adherence than standard care pelvic floor muscle exercises which can show low adherence–of the 17.774 women in the Norwegian Mother and Child Cohort Study, only 27.6% did pelvic floor muscle training at least three times a week at six months postpartum (Bø et al., 2007). Some women may see themselves as “not an exercise person” (Lagro-Janssen et al., 1994, translation by CO), find pelvic floor muscle exercises an inconvenience (Royal College of Midwives, 2018), or have an aversion to doing them (MacInnes, 2008). An additional effective method to strengthen the pelvic floor muscles would increase women’s choices in pelvic floor care. Although pelvic floor muscle training may technically be more effective and some women may be highly motivated to train (Mørkved and Bø, 2014), others could still prefer the balls, find them easier and more fun to use and adhere to, and be willing to trade off some training benefit, and the balls might therefore be more effective pragmatically.

On the other hand, it is important not to recommend, use or spend money on a (potentially) ineffective device simply because it is available and advertised on the health market. From a preliminary scientific literature search into the topic, a dearth of scientific evidence became clear. The profession of midwifery might not like to be cited on commercial advertising or a customer product information (as e.g. in FUN FACTORY, no date-a) and potentially abused to raise sales figures in such cases where there has been no research to test and confirm effectiveness claims.

Subjective anecdotal evidence on the balls’ effectiveness with the described method of use fed back to the author sounded reassuring but is, in comparison to research, a nonsystematic source of knowledge (Cluett and Bluff, 2006). Similarly, test participants in a consumer journal described their pelvic floor tissue as more firm after ball use (Test-Club-Bericht, 2013). However, in contemporary healthcare with its paradigm of evidence-based practice (Cluett and Bluff, 2006, Bhargava and Bhargava, 2007, *Consumers united for evidence-based health care*, 2011), scientific evidence is desired to decide on the use of interventions. The present research therefore set the first step towards a critical scientific evaluation of the popular claim of effectiveness (and safety) of this marketed device which to date is available for sale and advertised for the postpartum period without such evidence.

1.2 Midwifery and pelvic floor rehabilitation

The role of the midwife in pelvic floor rehabilitation is apparent at different levels of professional regulation. World Health Organization (WHO) (2010) guidance describes the

content of postpartum care from an international point of view. In terms of pelvic floor health, it states that a skilled birth attendant during early postpartum care should enquire about micturition, urinary incontinence and bowel function, that women "with involuntary leakage of a small volume of urine should be taught pelvic floor exercises" (World Health Organization, 2010, p. 46), and that women with persistent urinary or/and faecal incontinence should be referred for medical treatment. As midwives are core professionals in the provision of postpartum care, these recommendations fall within their scope of practice.

The International Confederation of Midwives (ICM) defines the midwife as a professional giving "the necessary support, care and advice during pregnancy, labour and the postpartum period", which "may extend to women's health" (International Confederation of Midwives, 2017, internationalmidwives.org/who-we-are/policy-and-practice/icm-international-definition-of-the-midwife/)–this consequently includes pelvic floor rehabilitation. However, the midwife's autonomous professional role mainly comprises the care of healthy women in a physiological process (International Confederation of Midwives, 2013). Health promotion, such as e.g. teaching women pelvic floor muscle exercises to prevent later pelvic floor problems, is part of this care (Dunkley, 2000, Bogle, 2006, Finlay, 2006), whereas the treatment of urinary incontinence (and other pelvic floor symptoms) falls within the scope of medical doctors and physiotherapists to whom the midwife has to refer (International Confederation of Midwives, 2013).

The above mentioned WHO document is based to a great extent on the UK National Institute for Health and Care Excellence (NICE) guidelines on routine postnatal[9] care (Demott et al., 2006) and therefore similar in content. A more recent UK paper, the *Joint statement on Pelvic Floor Muscle Exercise* by the Royal College of Midwives (RCM) and the Chartered Society of Physiotherapy (CSP), explicitly highlights the importance of the midwife in pelvic floor health promotion and the collaboration between midwives and physiotherapists (Gerrard and ten Hove, 2013).

To make pelvic floor muscle exercises a more acknowledged part of midwifery care has also been suggested by other authors (Mason et al., 2001b, Reilly et al., 2002). Guerrero et al. (2007), in a survey amongst women, midwives and obstetricians, came to the conclusion that midwives are best placed to and should approach the topic in the antenatal period. Since 2014, the RCM has been offering an i-learning module about anatomy, function and rehabilitation of the pelvic floor (Royal College of Midwives, no date). Finally, the topic of

[9] Although the term *postnatal* strictly seen refers to the newborn after birth (Farlex, 2018n), it is used in this thesis to refer to the mother after birth (who strictly seen is denoted by the term *postpartum*) as commonly used in the English language. The situation is similar with the term *antenatal* (Farlex, 2018a).

pelvic floor rehabilitation is contained within the professional midwifery literature (examples in English language being Dunkley, 2000, Nolan, 2001, Bick et al., 2004, Macdonald et al., 2011, Marshall et al., 2014, Medforth et al., 2017).

1.3 Childbearing, midwifery and pelvic floor rehabilitation in Austria

As the PhD topic arose out of Austrian midwifery practice and the empirical part of the project was performed in Austria, an overview on midwifery and pelvic floor muscle rehabilitation in Austria shall help understand the context of this work. As little data or literature is available on midwifery and routine pelvic floor care around childbirth in Austria, this section has been written from the author's Austrian midwifery experience, with the content verified in individual meetings with a midwife and a physiotherapist.

The Austrian Midwifery Act (Bundeskanzleramt Rechtsinformationssystem, 1994) ensures that every woman is cared for by a midwife during birth. Depending on her healthcare circumstances however, a woman might never see a midwife in pregnancy nor after birth. Usual antenatal care in Austria is delivered by obstetricians, and although every woman since 2012 has been entitled to one free midwifery consultation in pregnancy (Bundesministerium für Gesundheit, 2016), only about a quarter of pregnant women make use of this (Österreichisches Hebammengremium, 2018d). A woman might meet a midwife at the antenatal hospital care when registered to give birth in a hospital, where 98.4% of live births in Austria take place (Statistik Austria, 2018b). Only a woman opposed to this maternal healthcare system would organise complete midwifery care throughout pregnancy herself, which would disadvantage her of the public financial incentive of approximately 2000 Euro for the visits at the obstetrician (Geburtsallianz Österreich, no date).

Postpartum care is fragmented between nurses, midwives and obstetricians, with most women being cared for by nurses in the first few days after birth in hospital and all seeing their obstetrician for their six weeks postpartum check. Midwives only occasionally work on postnatal wards, and only a minority of women organises postpartum midwifery care at home (12% of postpartum women in 2012 (Österreichisches Hebammengremium, 2018c)). In hospitals, physiotherapists might routinely visit women during their stay on the postnatal ward. They give advice about the pelvic floor and its rehabilitation, including pelvic floor muscle exercises, and women might get a sheet with instructions for pelvic floor exercises in the postpartum period.

However, more than half of the midwives working outside of the hospital (over 900 of approximately 1700, according to Österreichisches Hebammengremium (2018b) and

Österreichisches Hebammengremium (2018a)), and also a number of those working exclusively in the hospital offer antenatal childbirth education classes which can be assumed, as this is part of midwifery care, to include the topic of pelvic floor muscle rehabilitation. Nearly half (47.6%) of all Austrian midwives offer home visits after childbirth (Österreichisches Hebammengremium, 2018b). For the later post partum period, 398 of the 1708 midwives working outside of the hospital (23.3%) offer postpartum rehabilitation classes with exercises for abdominal, back and pelvic floor muscles (as per 22 July 2018, Österreichisches Hebammengremium, 2018b).

The only quality standard with respect to peripartum pelvic floor care is a guideline on 3rd and 4th degree perineal tears during childbirth from the Austrian Urogynaecology Working Group (Arbeitsgemeinschaft für Urogynäkologie und rekonstruktive Beckenbodenchirurgie Österreich, 2014). Pan-Austrian information material on pelvic floor muscle rehabilitation in childbearing has been created only recently (Medizinische Kontinenzgesellschaft Österreich, 2016)[10]. It cannot be ascertained whether the health professionals involved in care around childbearing routinely recommend pelvic floor muscle exercises and antenatal and/or postnatal classes to all pregnant and postpartum women. The only statistical information on the number of women attending antenatal classes comes from the clientele of five public hospitals in Vienna where it has been found to be 21%, with the women using these services having higher education and living in better social circumstances (Wimmer-Puchinger et al.). In any case, class attendance is not defrayed by the public health insurance system and women have to pay for themselves.

1.4 Thesis features

To gain an overview on the following work, research interest and design of this PhD project and the structure of the thesis are explained.

1.4.1 Research interest and design

This thesis arose out of a research interest into the effectiveness of vibrating vaginal balls in the childbearing time; more specifically on their effectiveness *after* childbirth because the balls are usually meant to rehabilitate the pelvic floor after pregnancy and birth. Thereby, the method of use in question is to keep the ball in the vagina while moving around and to utilise its weight and vibration effect. According to the midwife's professional role, this original

[10] An algorithm for pelvic floor care of postpartum women is currently being tested at different Austrian hospitals (Udier, 2018). –Elisabeth Udier is an Austrian physiotherapist specialised in women's health physiotherapy. She is vice-president of the professional network for uro-, procto- and gynaecology and obstetrics of the Austrian physiotherapy association Physio Austria.

research interest focuses on ball use to enhance pelvic floor muscle performance to prevent later pelvic floor impairments (and not on pelvic floor treatment of any kind).

The doctoral journey started with a systematic review which showed that further research on the effectiveness of vibrating vaginal balls after childbirth is warranted. The best research design to test the effectiveness of an intervention is a randomised controlled trial (RCT) (OCEBM Levels of Evidence Working Group, 2011). In an RCT, the outcomes of groups are compared, with one group using an experimental intervention and the other(s) using a comparison or no intervention (Meinert, 2012); the participants are allocated to the groups at random. With variables that could also influence the outcome being controlled for by different techniques, the intervention as the factor being varied across groups allows to conclude on its causal influence (Hulley et al., 2013).

RCTs need careful planning and preparation of all methodological components so that they can run together smoothly during the trial. One way of ensuring the success of an RCT is to perform a preliminary feasibility trial to try out the different methodological aspects. Resulting from the lessons learnt, necessary adjustments to the original research plan can then be made to optimise a main trial, or it can be seen that the trial is not feasible as planned (Thabane et al., 2010). Feasibility studies are encouraged and funded by funding bodies (e.g. the National Institutes of Health, 2016, National Institute for Health Research, 2017, Health Research Council of New Zealand, no date).

Therefore, and in view of the available time and resources, the empirical PhD project was designed as a feasibility trial to test the feasibility of a future full RCT on the effectiveness of vibrating vaginal balls to improve pelvic floor muscle performance in women after childbirth. As such, this trial set out to test recruitment and sampling procedures, application and safety of the experimental and comparison interventions, adherence to the interventions, pelvic floor muscle performance measurement, data collection and analysis. It should further provide preliminary scientific evidence on clinical results. Alongside these issues, women's experiences with and perspectives on the interventions and the trial were of interest.

1.4.2 Overview of thesis chapters

This thesis is organised into 11 chapters, as follows:

This introduction chapter (Chapter 1) is followed by two background chapters. The first of these, Chapter 2, provides knowledge on female pelvic floor morphological and functional anatomy, and dysfunction as relevant for this thesis. It further informs on pelvic floor changes and pelvic floor muscle rehabilitation in the childbearing period. In Chapter 3 follows

a review of knowledge on the use of vaginal cones or balls for pelvic floor muscle rehabilitation with a focus on the postpartum time. This knowledge is complemented by a systematic review on the effectiveness of vibrating vaginal balls or similar devices to improve pelvic floor muscle performance and urinary continence in women after childbirth.

The thesis purpose and design chapter (Chapter 4) gives the rationale for the empirical research with its aims, questions and objectives. It brings an overview of the research design adopted in this study by presenting the general features of an RCT and of feasibility research. The methods chapter (Chapter 5) sets out in detail the methodology implemented in the feasibility trial. Its sections on feasibility methodology, recruitment and sampling, the interventions, data collection and analysis, ethical issues, and patient and public involvement (PPI) describe and justify the way in which the trial was conducted and the decisions made with respect to these areas, and how rigour was ensured.

Chapters 6 to 9 present the study findings in terms of trial processes (including women's experiences with and opinion on the interventions and the trial), resources and management outcomes and clinical results. Each of these chapters also discusses and interprets the presented results and refers to relevant literature in order to try to explain the trial's findings and to put the results in context.

Chapter 10 provides a synthesis of the different aspects of the trial findings with reference to the overarching research question on trial feasibility posed in this thesis. From this, it points out the conclusions with respect to potential future research, in particular a full RCT. It names possible implications for trial design, sample size, processes, management and resources. The importance of PPI work for research decisions is stressed.

The conclusions chapter (Chapter 11) highlights the new knowledge generated by the systematic review and feasibility RCT. The feasibility trial's strengths and limitations are addressed. The chapter finally indicates the significance of the thesis results and their contribution to the larger body of knowledge. It does so by summarising the implications that can be drawn from the findings with respect to further research and by delineating recommendations for the use of vibrating vaginal balls to strengthen the pelvic floor muscles after childbirth.

Where applicable, the trial reporting corresponds to the CONSORT (Consolidated Standards of Reporting Trials) guidelines for feasibility and pilot studies (Eldridge et al., 2016a), complemented by the CONSORT extensions for nonpharmacologic treatments (Boutron et al., 2008), patient-reported outcomes (Calvert et al., 2013), harm (Ioannidis et al., 2004), and

abstracts (Hopewell et al., 2008); and to guidelines for reporting interventions (Hoffmann et al., 2014), PPI (Staniszewska et al., 2017), and qualitative research (Tong et al., 2007). Reporting of statistical results considered the SAMPL guidelines (Lang and Altman, 2014), the American Psychological Association (2010) *Publication Manual*, Petrie and Sabin (2009), and Thabane and Akhtar-Danesh (2008).

1.5 Research team and venue

The trial could not have been performed had there not been the support of a team of collaborators for the principal investigator (PI) Claudia Oblasser (CO). All nine team members are named in Appendix B with a description of their contribution to the project. Complementary to City University London[11] being the primary sponsor, the Medical University of Vienna (Medizinische Universität Wien) and Vienna General Hospital (Allgemeines Krankenhaus der Stadt Wien, AKH Vienna) as its clinical centre provided support in terms of room, material and working hours of Engelbert Hanzal (EH).

Summary

This chapter has delineated the research problem and relevance of the topic. It has outlined the role of the midwife in pelvic floor care on an international level and presented an overview on childbearing, midwifery and pelvic floor care in Austria. The research interest and design were pointed out: This study, with its empirical part designed as a feasibility trial, was interested in preparing a future RCT on the effectiveness of vibrating vaginal balls after childbirth to improve pelvic floor muscle performance. An overview on the structure of the thesis by a summary of its chapters was given. The research team and venue were presented. To help understand the physiology of pelvic floor health, the next chapter provides background knowledge on the female pelvic floor and childbearing.

[11] City University London became City, University of London in September 2016. The former name is used in this thesis when referring to events that took place under the former name.

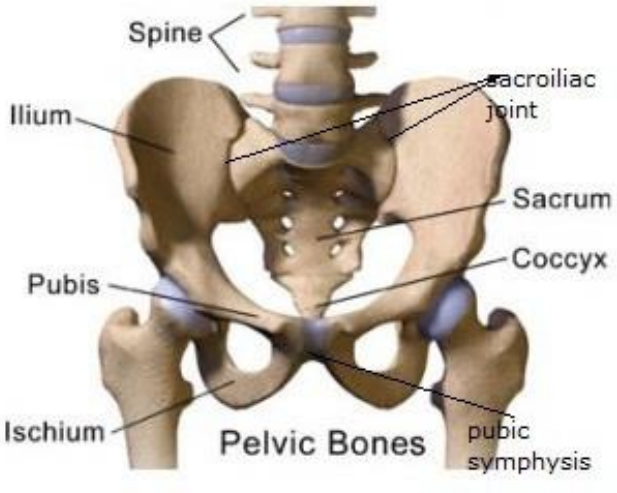
2 THE PELVIC FLOOR AND CHILDBEARING

This first background chapter shall help to gain an understanding of the pelvic floor and its changes by childbearing. It first describes the morphological anatomy of the pelvic floor, summarises its roles and explains the parts of its functional anatomy which are of relevance to understand the research context of this PhD project. During childbearing, the pelvic floor is exposed to various changes and challenges which are described in the following section. Thereafter, a closer look is taken at pelvic floor muscle rehabilitation during the childbearing period in general and with a focus on preventive rehabilitation post partum.

2.1 Morphological anatomy of the pelvic floor

The pelvis forms the lower part of the human trunk. Figure 2 shows its bones with their connections, and in a general sense, the pelvic floor is the muscular closure of the pelvic outlet (Pschyrembel, 2013). From an in depth anatomical point of view however, and this definition is adopted in this thesis, the term pelvic floor only denotes the pelvic diaphragm, the superior (cranial) layer of the muscular closure with its fasciae, with the inferior (caudal) layer being called the perineum (Federative International Programme on Anatomical Terminologies, 2011, Paulsen and Waschke, 2013, Marieb, 2015).

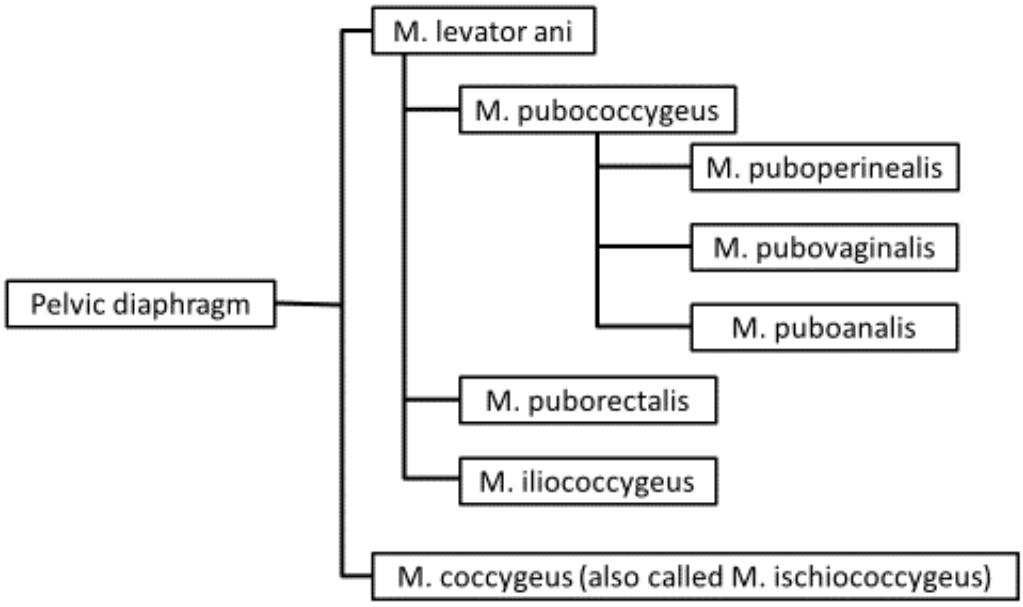
Figure 2 The bony pelvis (generic)



Source: IncoStress (2006). Item reprinted with permission of CEO C&G Medicare Ltd.

Figures 4-7 depict the pelvic diaphragm; Figure 3 lists its muscles according to the international standard on human anatomic terminology (Federative International Programme on Anatomical Terminologies, 2011). The muscles consist of left and right instances: The levator ani, the largest muscle, takes its origin anteriorly from the pubic bone and laterally from the tendinous levator arch (arcus tendineus musculi levatoris ani, ATLA) and the ischial spine (Spina ischiadica in Figures 4 and 6); it inserts medially to the walls of the vagina, rectum and anal canal, the perineal and anococcygeal body (fibromuscular tissue structures), and the coccyx (Dykes, 2003, Marieb, 2015). The opening between the lateral instances of the levator ani is called the (urogenital and anal) levator hiatus (Köpf-Maier and Wolf-Heidegger, 2000, Delancey, 2016). It serves as a passageway for the urethra, vagina and anus–the outlets of bladder, uterus and rectum, the pelvic organs lying above the pelvic diaphragm (Fritsch, 2012).

Figure 3 Tree of the muscles of the pelvic diaphragm according to the Federative International Programme on Anatomical Terminologies (2011)



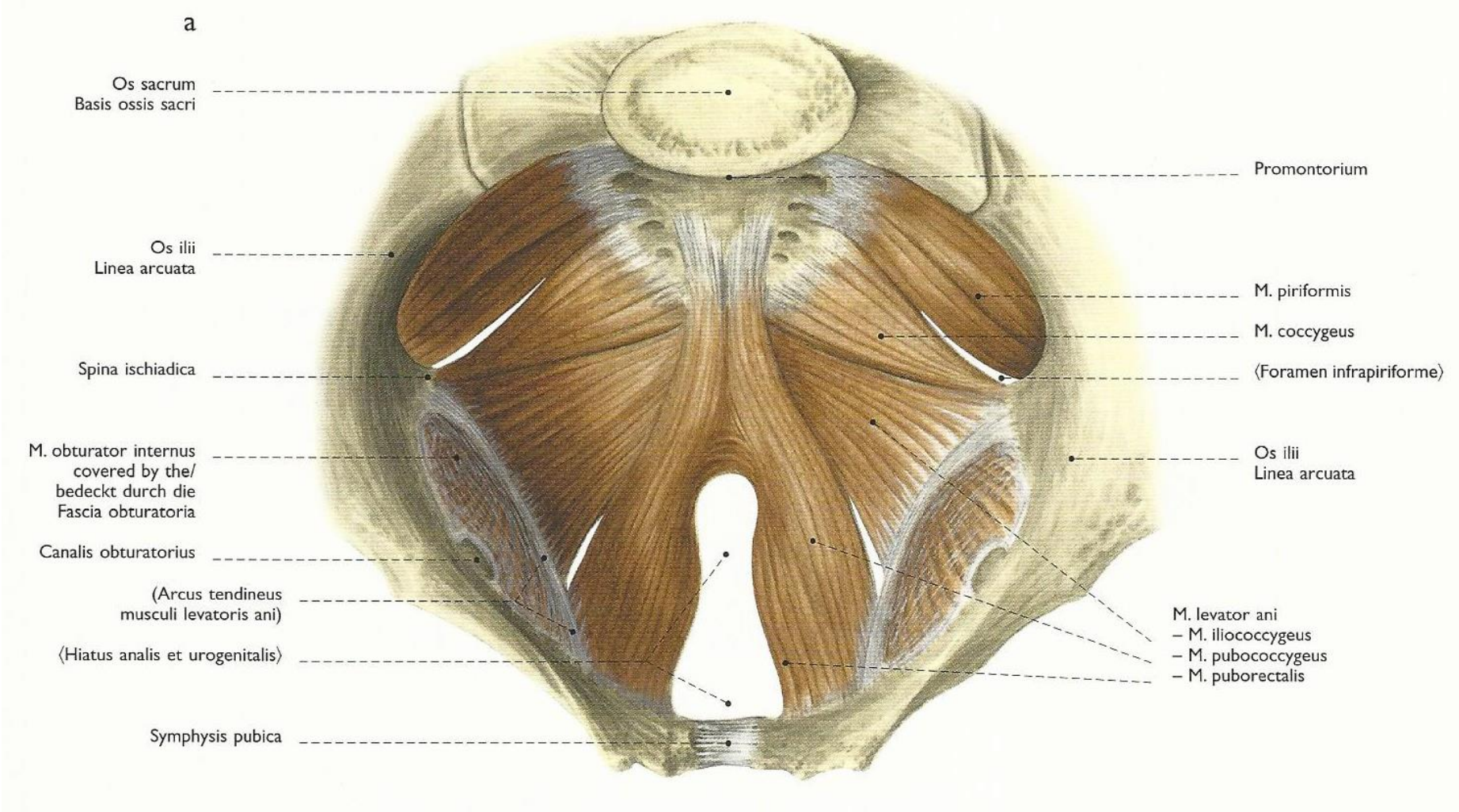
M. = muscuclus (muscle).

Inferior to the pelvic diaphragm lies the perineum[12] (Marieb, 2015, Figure 7) which is divided into the urogenital and anal triangles (Federative International Programme on Anatomical Terminologies, 2011, Delancey, 2016). In terms of muscular structures, the anal triangle contains the external anal sphincter. The urogenital triangle consists of the deep layer muscles deep transverse perineus and rhabdosphincter (also called striated urogenital or external urethral sphincter muscle, with its parts sphincter urethrae, compressor urethrae and sphincter urethrovaginalis), and the superficial layer muscles superficial transverse perineus, bulbospongiosus, and ischiocavernosus (Perucchini and DeLancey, 2008, Tortora and Derrickson, 2011, Delancey, 2016, Fry et al., 2017).

The pelvic muscles and organs are surrounded and held together by various connective tissue structures in the form of fasciae and ligaments (e.g. the above mentioned ATLA) (Delancey, 2016). The muscles are innervated by neural structures stemming from the sacral nerve plexus (formed by lumbar and sacral spinal nerves (Pschyrembel, 2013)) and the pudendal nerve (Vodušek, 2008). The levator ani is innervated by the levator ani nerve with a possible contribution by the pudendal nerve (Barber et al., 2002, Shobeiri et al., 2008, Wallner et al., 2008).

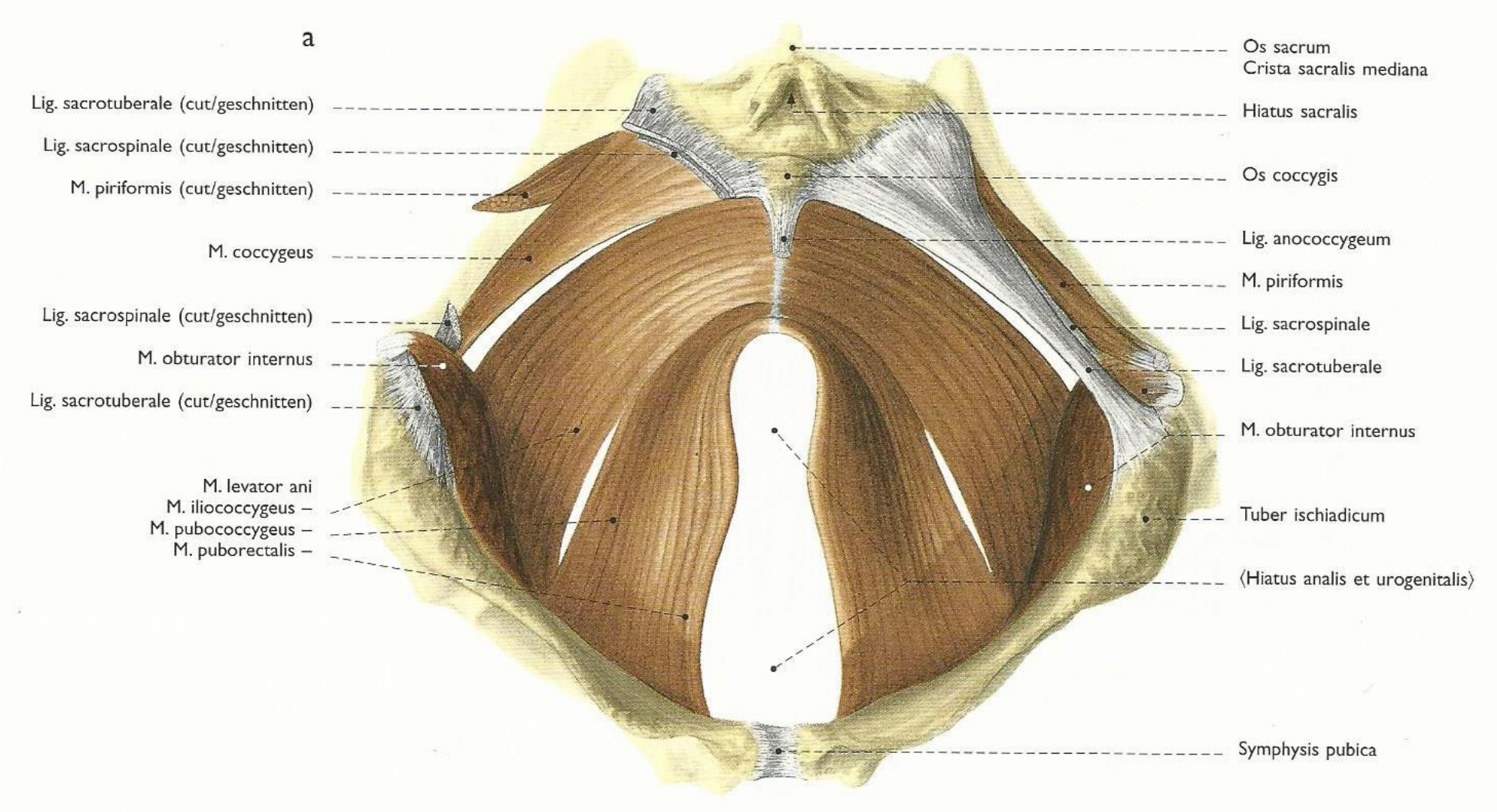
[12] The term perineum here denotes the *anatomical* perineum, the area extending from the pubic bone to the coccyx, with underlying tissues (Federative International Programme on Anatomical Terminologies, 2011, Tiran, 2012, Delancey, 2016). The *obstetric* perineum in contrast is equated with the perineal body (centrum perinei) which is the fibromuscular centre of the perineum (Federative International Programme on Anatomical Terminologies, 2011, Tiran, 2012).

Figure 4 The female pelvic diaphragm, superior aspect



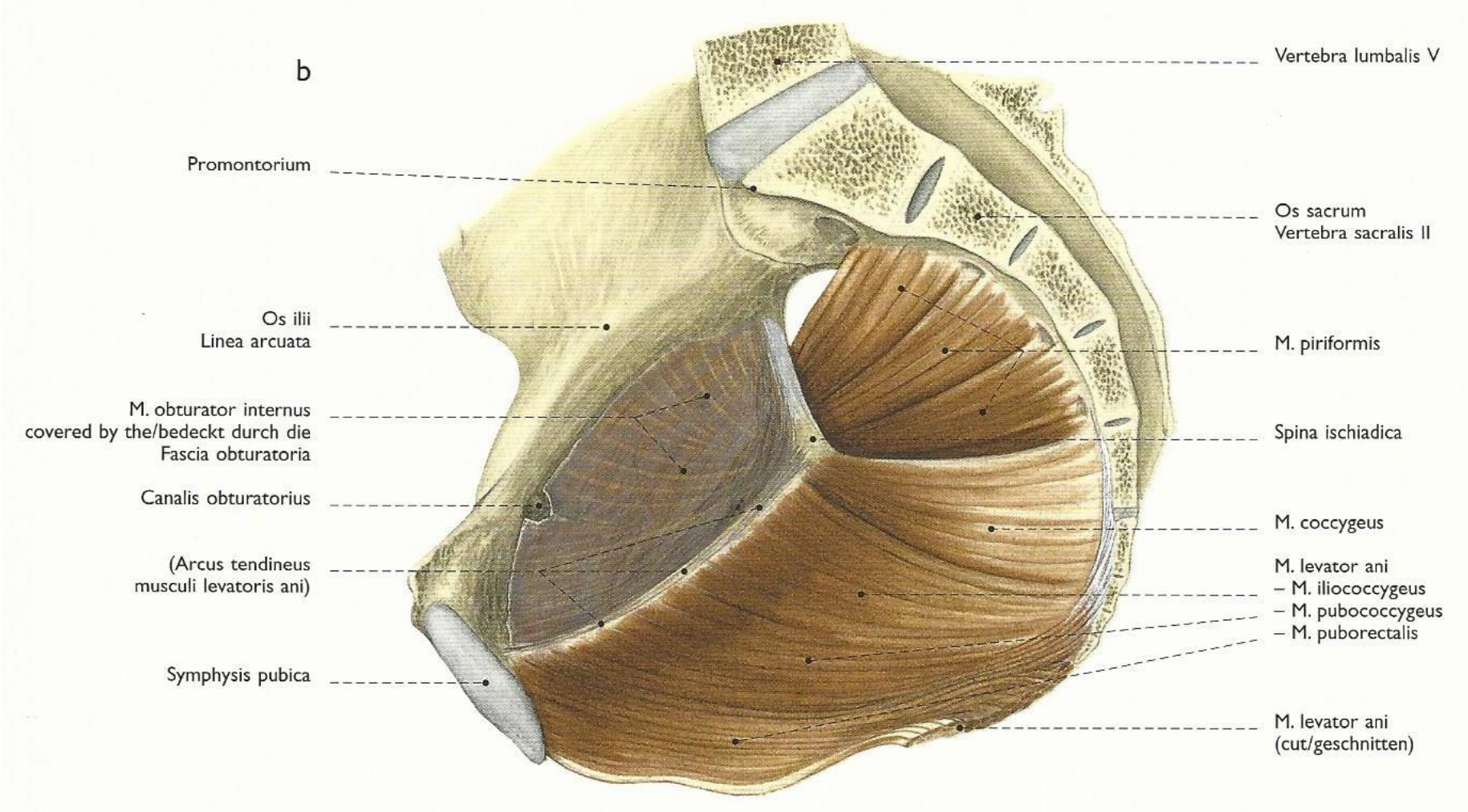
Source: Köpf-Maier and Wolf-Heidegger (2000). M. = muscuclus (muscle).

Figure 5 The female pelvic diaphragm, inferior aspect



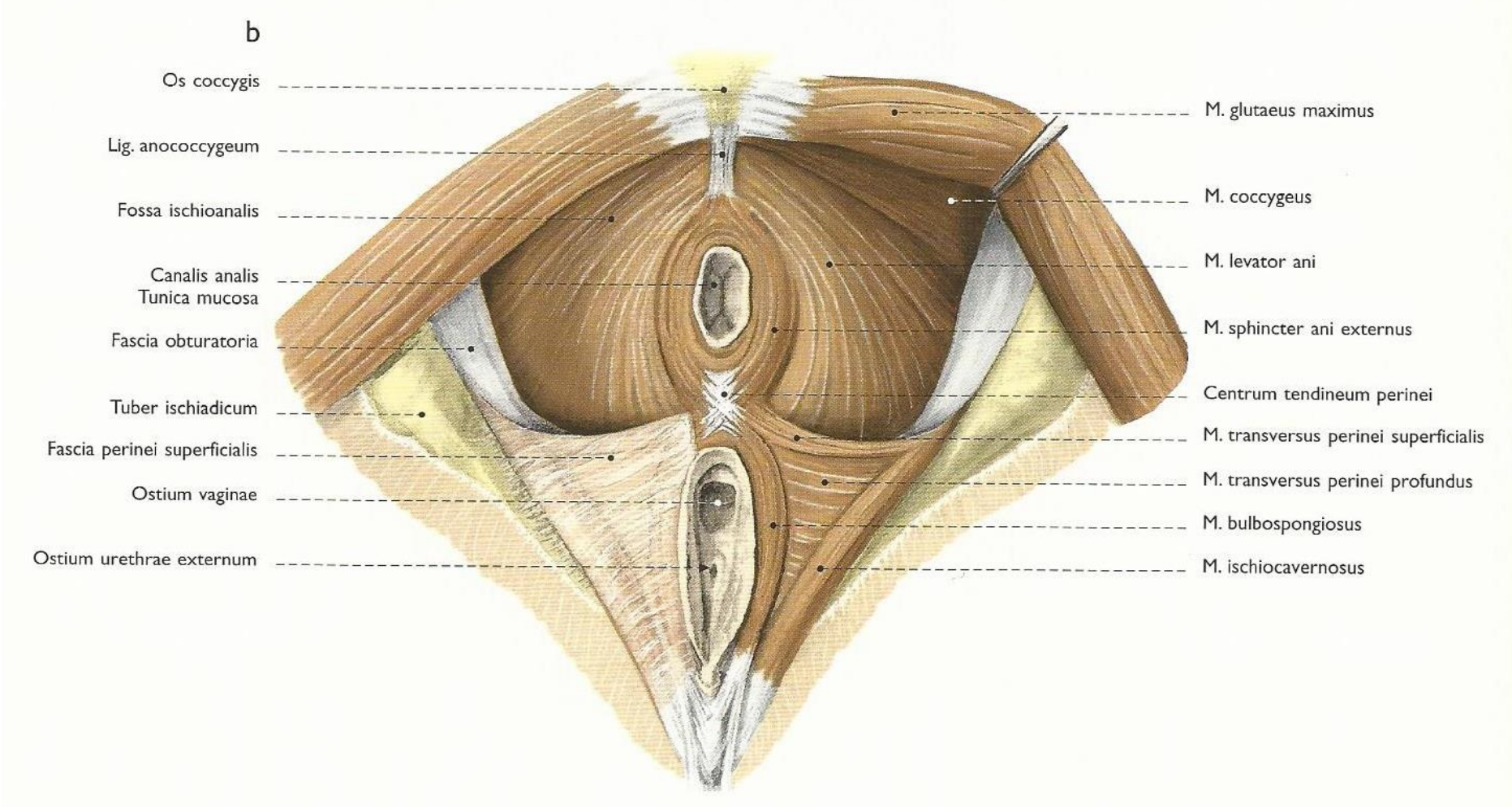
Source: Köpf-Maier and Wolf-Heidegger (2000). M. = muscuclus (muscle); Lig. = ligamentum (ligament).

Figure 6 The female pelvic diaphragm, medial aspect of the right half



Source: Köpf-Maier and Wolf-Heidegger (2000). In slight contrast to this usual basin shaped drawing, Shafik (1979) and Hjartardottir et al. (1997) have shown that the pelvic floor is funnel (from a medial view, or “dome” when considering each lateral half) shaped, i.e. the most medial lower aspect of the drawing should be formed like the inner half of a funnel. M. = muscuclus (muscle).

Figure 7 Muscles of the female pelvic diaphragm and perineum, inferior view



Source: Köpf-Maier and Wolf-Heidegger (2000). M. = muscuclus (muscle); Lig. = ligamentum (ligament).

2.2 Functional anatomy and dysfunction of the pelvic floor, including the importance of proper pelvic floor muscle performance

The pelvic floor fulfils various functions. It closes the lower opening of the pelvic cavity and provides structural support to pelvic stability and the pelvic organs (Gödl-Purrer, 2006, Heller, 2015). It contributes to micturition and defaecation and helps to maintain urinary and faecal continence (Delancey, 2016). It influences sexual function and allows and facilitates childbirth (Brayshaw, 2003, Coad and Dunstall, 2011). Finally, it serves as venous and lymphatic pump for the pelvis (Beyond Basics Physical Therapy, 2018) and supports postural[13,14] and respiratory functions (Hodges et al., 2007).

Pelvic floor function depends on the integrity of its muscles, fascia and nerves (Bø et al., 2015). The pelvic floor muscles are voluntary skeletal muscles and their activity consists of two types: tonic and phasic (Vodušek, 2008). Tonic means that the pelvic floor muscles as postural muscles have a higher resting tone than other skeletal muscles due to their task of maintaining continuous contraction except during micturition, defaecation and the Valsalva[15] manoeuvre (Schwertner-Tiepelmann et al., 2012). Phasic denotes the muscle activation by reflectory[16] tension during rises in intraabdominal pressure and by voluntary muscle contraction (Sapsford and Hodges, 2001, Junginger et al., 2010). For the two activity types, the pelvic floor muscles consist of two types of muscle fibres (Perucchini and DeLancey, 2008). Type I slow-twitch muscle fibres are high in endurance and responsible for the muscular tone at rest, type II fast-twitch muscle fibers are strength/power[17] fibres and required for reflex[18]-related and voluntary contractions (Perucchini and DeLancey, 2008, Ratamess, 2012).

Functional anatomy areas with associated dysfunctional disorders of the pelvic floor, as far as relevant for this thesis, are explained in the following sections. This includes the contribution of pelvic floor muscle performance to continence, structural support of the pelvic organs and sexual function.

---

[13] Contributing to body posture (Farlex, 2018j).
[14] Please note: Uncommonly used medical terms and methodological scientific terms are collected in the Glossary in Appendix JJ if they appear in the text more than once and in different sections.
[15] "[A]ny forced expiratory effort against a closed airway, such as when an individual holds the breath and tightens the muscles in a concerted, strenuous effort to move a heavy object" ('Valsalva maneuver', 2009, htpp://medical-dictionary.thefreedictionary.com/Valsalva+manoever).
[16] "Designating or relating to an action […] caused by an automatic response to a stimulus" (Oxford University Press, 2018); also called reflexive (Farlex, 2018l).
[17] Power: product of strength and speed of movement (Wilmore and Costill, 2004).
[18] "[A]utomatic, involuntary response to a stimulus" ('Reflex', 2014, https://www.thefreedictionary.com/reflex).

Urinary continence

The anatomical urinary continence system consists of two subsystems: the sphincteric closure system and the urethral support system (Ashton-Miller et al., 2001). The sphincteric closure system consists of the striated urogenital sphincter muscle (see section 2.1) and a smooth muscle sphincter (Delancey, 2016). The pelvic floor is seen as the urethral support system (DeLancey, 1990, Ashton-Miller and DeLancey, 2015), and the relevance of its function to urinary continence is explained by DeLancey (1990) with a theoretical analogy: The pelvic floor works like a surface under a garden hose (the urethra)–if the surface is noncompliant, stepping on the hose stops the water (urine) flow; if it is compliant, stepping on the hose does not stop the water flow. To guarantee continence during e.g. coughing, sneezing, or physical activity, the pelvic floor muscles must be able to contract reflexively[19], rapidly, and strongly (Luginbuehl et al., 2012).

The dysfunctional symptom of urinary incontinence comes in different types. Among other types, stress urinary incontinence is defined as the "complaint of involuntary loss of urine on effort or physical exertion (e.g. sporting activities), or on sneezing or coughing", urge incontinence as the "complaint of involuntary loss of urine associated with urgency", and mixed urinary incontinence as a combination of the two (Haylen et al., 2010, p. 7).

In accordance with the hose theory, a number of studies found that the pelvic floor muscles of urinary continent women performed better than those of incontinent women (Sampselle, 1990, Hahn et al., 1996, Bø, 2003, Mørkved et al., 2004, Amaro et al., 2005, Thompson et al., 2006, Baracho et al., 2012, Hilde et al., 2012, Hilde et al., 2013b, Luginbuehl et al., 2015a). Accordingly, Luginbuehl et al.'s (2015a) review confirmed an association between improvement of pelvic floor muscle performance and a reduction in symptoms after physical therapy for urinary incontinence. Likewise, subjective contraction strength after childbirth is lower in women with stress urinary incontinence (Dietz et al., 2012). Further, the difference between contraction and relaxation muscle thickness of the muscular layer caudal to the pelvic diaphragm and anterior to the anorectum (urogenital diaphragm) was higher in continent women (Mørkved et al., 2004), and the internal levator ani surface area at rest was reduced after training, suggesting a higher muscle tone (Dumoulin et al., 2007).

Anal continence

Anal continence is maintained by the external and internal anal sphincters together with support from the pelvic floor (Salvatore et al., 2017). The levator ani's puborectalis muscle

[19] See footnote 16.

sling (Figures 4-6) pulls the rectum forward to create the anorectal angle and rectal valve, which both constitute closure mechanisms (Jung et al., 2007, Quigley, 2007, Sultan and Abulaffi, 2008). Pelvic floor muscle weakness may thus be a factor contributing to anal incontinence (Berghmans and Bols, 2015, Salvatore et al., 2017). Anal incontinence comprises flatus (involuntary loss of wind) or faecal (involuntary loss of solid or liquid faeces) incontinence (Haylen et al., 2010).

Organ support

The pelvic organs (bladder, urethra, uterus, vagina, rectum) are supported and held in place by the pelvic floor muscles and connective tissue structures (Delancey, 1993). A pelvic organ support system deficiency results in pelvic organ prolapse, clinically seen as descent of the anterior or/and posterior vaginal wall (cystocele/rectocele) or/and the cervix/uterus through the urogenital hiatus (Haylen et al., 2010). Symptoms consist in a “departure from normal sensation, structure, or function, experienced by the woman in reference to the position of her pelvic organs” and includes vaginal bulging or/and pelvic pressure, heaviness or dragging (Haylen et al., 2010, p. 6). Suboptimal pelvic floor muscle performance has been shown to be a factor contributing to pelvic organ prolapse (Bø and Frawley, 2015, Salvatore et al., 2017), and subjective contraction strength after childbirth is lower in women with pelvic organ prolapse symptoms (Dietz et al., 2012).

Sexual function

The levator ani is ascribed a role for sexual function (Kegel, 1952, Graber and Kline-Graber, 1979). While it is assumed that stronger pelvic floor muscles enhance sexual function (McKey and Dougherty, 1986, Lowenstein et al., 2010, Yeniel and Petri, 2014), Martinez et al. (2014) concluded that any cause-effect relation between strong pelvic floor muscles and sexual function and its direction is yet to be postulated. Among sexual dysfunction complaints which might be ascribed to impaired pelvic floor muscles are vaginal laxity (Haylen et al., 2010) and arousal and orgasmic disorders (Basson et al., 2001). In Tennfjord et al. (2015), postpartum participants reporting vaginal looseness or laxity had lower vaginal manometry values than women without this symptom. Indirectly, pelvic floor performance influences sexual function in that conditions related to pelvic floor dysfunction (e.g. pelvic organ prolapse and urinary and anal incontinence symptoms) correlate with sexual dysfunction (Handa et al., 2007, Rosenbaum, 2007, Handa et al., 2008).

2.3 Changes and challenges of the pelvic floor in childbearing

Changes and challenges of the pelvic floor in each of the three phases of childbearing–pregnancy, birth and puerperium–are described in the following sections, focusing on pelvic floor muscle performance.

2.3.1 Pregnancy

From the beginning of pregnancy, hormonal changes serve to maintain the pregnancy and prepare the female body for childbirth. The corpus luteum and placenta produce (amongst others) the pregnancy related hormones oestrogen, progesterone and relaxin (Costanzo, 2010, Sherwood, 2016), with progesterone stimulating the production of the enzyme collagenase (Silbernagl and Despopoulos, 2012). Supported by the increased blood flow to the pelvic region (Van Rooyen, 1969), these biochemical changes lead to structural and functional changes in the pelvic floor (Chaliha, 2006). The biochemical influences are fortified by the mechanical changes of the growing uterus requiring space as well as by the physiological weight gain (Baessler and Schüssler, 2008, Sangsawang and Sangsawang, 2013).

The connective tissue becomes more elastic (Van Rooyen, 1969, Landon et al., 1990, Lavin et al., 1997). This is thought to contribute to urinary incontinence, although in Kristiansson et al. (2001) a higher relaxin level unexpectedly correlated with a lower rate of stress urinary incontinence developing during pregnancy. Hiatal dimensions and perineal body length increase (O'Boyle et al., 2003, Shek et al., 2012b). Wijma et al. (2001) showed a lowering of the pelvic floor and a decrease in reflectory pelvic floor muscle contraction in pregnancy, and Peschers et al. (1996) as well as Shek et al. (2012b) demonstrated increased urethral mobility[20]. Caroci et al. (2014) classified pelvic floor muscle strength in the first trimester of pregnancy as weak in the majority of participants although they stressed that there is no consensus on adequate pelvic floor muscle strength. Palmezoni et al. (2017) measured a reduced pelvic floor muscle strength in primiparous pregnant women compared to nulliparae.

Together with pregnancy-induced changes in the urinary tract, these pelvic floor changes lead to the common urinary symptoms in pregnancy–nocturia[21], increased urinary frequency and voiding difficulties (Chaliha, 2006). Between 2% and 50%[22] of pregnant women report any urinary incontinence, and 7.4% to 85% report stress urinary incontinence, the type of

[20] Urethral/bladder neck mobility in the non-pregnant healthy woman is low.
[21] Nocturnal urination (Farlex, 2018h).
[22] Variation in urinary incontinence rates may be explained by methodological differences–different populations investigated, use of different definitions of incontinence and registration of incontinence at different stages of pregnancy or after birth (Mørkved, 2007).

urinary incontinence most commonly associated with pregnancy (Viktrup, 2002, Sangsawang and Sangsawang, 2013). De novo, urinary stress incontinence appeared in 29% to 39.1% (in primiparous 50.0%) of pregnant women (Viktrup et al., 1992, Marshall et al., 1998, Solans-Domenech et al., 2010). Overall, incontinence prevalence increases during pregnancy (Viktrup et al., 1992, Thorp et al., 1999, Kristiansson et al., 2001).

Flatal incontinence was reported by 34.6% of women by the 12th week of the first pregnancy, and by 42.3% at 36 weeks gestation, faecal incontinence by 3.9% and 3.0% respectively (van Brummen et al., 2006). The prevalence of pelvic organ prolapse has been shown to increase with and during pregnancy (Baessler and Schüssler, 2008), up to a rate of 48% in the second or third trimester (O'Boyle et al., 2002). In their systematic review on sexual function in childbearing, Yeniel and Petri (2014) found that in three of four studies screened, the ability to orgasm declined throughout pregnancy. During sexual activity, vaginal contractions at climax are weaker in the third trimester, and tonic muscle spasms do sometimes occur (von Sydow, 1999). Erol et al. (2007) reported orgasmic disorder in 81% of healthy participants.

2.3.2 Vaginal birth

The passage of the baby through the vagina is a stressful event for the pelvic floor. Svabik et al. (2009), assuming an optimal fetal position in a Caucasian population, calculated the levator hiatus having to distend between 25% and 245% from dimensions at maximal Valsalva. In Lien et al.'s (2004) biomechanical computer simulation on the stretching of the levator ani muscles during vaginal childbirth, the largest tissue strain reached a stretch ratio of 3.26 in the medial pubococcygeus muscle, with regions of the ileococcygeus and puborectalis muscles reaching maximal stretch ratios of 2.73 and 2.28. Similarly, Martins et al. (2007) showed maximum pelvic floor muscle stretch during childbirth to be 1.6. Stretching of the pudendal nerve during childbirth has been modelled to reach a maximum of 35% to 41% by Lien et al. (2005).

While genital birth injuries can be visible, trauma to the pelvic floor tissue may also occur without obvious injuries (Shek et al., 2011). Apart from stretching with ensuing muscle rupture and/or denervation, compression with ischaemia[23] and a resulting reduced oxygen supply may occur (Baessler and Schüssler, 2008, DeLancey and Ashton-Miller, 2015). The (over)stretching and/or ischaemia can induce mechanical and/or biochemical trauma to the pelvic floor tissue structures (Dietz and Schierlitz, 2005, DeLancey et al., 2007, Ashton-Miller

[23] Insufficient blood flow (Farlex, 2018f).

and Delancey, 2009). Birth trauma can result in postnatal symptoms as described in the next section.

2.3.3 Post partum

Women after childbirth can suffer from pelvic floor symptoms, but the postpartum woman's pelvic floor has also been researched by various physiological measurements. Thus, symptoms and measurements are considered in the next two sections for the earlier postpartum time, whereas the third section looks at these outcomes in the longer term.

Pelvic floor symptoms

Compared to pregnancy, the prevalence of urinary incontinence following delivery seems to decrease, although postpartum prevalence remains higher than before pregnancy (Stanton et al., 1980, Viktrup et al., 1992, Foldspang et al., 1999, Mason et al., 1999, Thorp et al., 1999, Viktrup and Lose, 2000). During the first three months after birth, 33% (95% confidence interval [CI] [32, 36], pooled value) of women are reported to experience urinary incontinence, with small changes over the first postpartum year (systematic review by Thom and Rortveit, 2010). Pooled estimates of any de novo urinary incontinence between 2 and 13 weeks post partum for primiparous women was 17.4% (95% CI [16.1, 18.7]), and 2.9% (95% CI [1.1, 4.6]) for stress incontinence.

De novo faecal incontinence six to seven months after childbirth was shown to be 4% by MacArthur et al. (1997). One year after the first birth, it was 2.6% in van Brummen et al. (2006), who also showed de novo flatal incontinence to be 8.5% at 12 months post partum. In a population-based survey (21,824 recipients, response rate 40%) by Guise et al. (2007), 29% of participating women within three to six months post partum reported experiencing anal incontinence since delivery, 46% of them incontinence of stool and 38% incontinence of flatus only; approximately 46% of women with anal incontinence reported onset of incontinence after delivery of their first child. Women can also suffer from vaginal flatus when, with changes of body position or during sexual activities, air leaving the vagina involuntarily leads to flatus sounds which cannot be suppressed (Neels et al., 2017). This flatus vaginalis is supposed to be caused by an incomplete closure of the vaginal introitus and insufficient support of the vaginal wall.

In 101 women, Sze et al. (2002) found a prevalence of pelvic organ prolapse of 83% at six weeks post partum, in 52% of all participants of stage II[24], in 37% de novo, in 15% more severe than in pregnancy. In the three to six months after birth, pelvic organ prolapse stage

[24] Pelvic organ prolapse is staged 0-IV (Haylen et al., 2010).

≥II was found to have a prevalence between 7.7% and 56% (Handa et al., 2009, Diez-Itza et al., 2011a, b, Wai et al., 2011).

Women after childbirth may experience a range of sexual difficulties (Brown and Lumley, 1998, Barrett et al., 2000, Abdool et al., 2009b, Leeman and Rogers, 2012, Yeniel and Petri, 2014, McDonald et al., 2015), of which some might be attributed at least partly to changes of pelvic floor muscle performance. In a questionnaire survey within the first year post partum by Jacobson et al. (1967), 19.9% of women reported feeling muscular hypotonicity in the abdominal wall and “the perineum”. In Barrett et al.’s (2000) study, women up to six months post partum reported vaginal looseness and lack of muscle tone compared with the year prior to pregnancy. In Baxter (1974), 14 of 48 women felt that, compared with before pregnancy, their vagina had been slacker on resumption of intercourse (at mean 6.2 weeks post partum), and 17 said it had felt tighter. By the time of the interview (between 11 and 15 weeks post partum) it was slacker for eight and tighter for 10 women. Yeniel and Petri (2014) summarised that reaching orgasm for women post partum was more difficult than before or during pregnancy, with recovery over the postpartum time. According to von Sydow (1999), orgasm is less intensive in the first six to eight weeks post partum and during breastfeeding.

Physiological measurements of pelvic floor characteristics

After vaginal childbirth, the hiatus is widened (Shek and Dietz, 2009). Peschers et al. (1996) showed a lower bladder neck, increased bladder neck mobility and decreased bladder neck elevation (task of the pelvic floor) after vaginal delivery. Vaginal childbirth is strongly associated with increased anterior vaginal wall descent (Dietz et al., 2002), and pregnancy and birth seem to be associated with an increase in prevalence and size of true rectoceles (Dietz and Steensma, 2006). The postpartum prevalence of traction neuropathy of the pudendal nerve has been shown by electrophysiological studies to be as high as 42% (Snooks et al., 1984). In comparisons between different groups, a reduced pelvic floor muscle performance after vaginal delivery was found (Marshall et al., 2002, Baytur et al., 2007, Sigurdardottir et al., 2011, Hilde et al., 2013b). Dietz et al. (2012) demonstrated that subjecive pelvic floor muscle contraction strength is affected by childbirth.

In up to 13-36% of women, and up to 65% in high risk groups such as women with forceps delivery, partial or full avulsion of the puborectalis muscle after childbirth has been documented (Dietz et al., 2012, Schwertner-Tiepelmann et al., 2012). Not all women notice the avulsion (Dietz et al., 2012, Thibault-Gagnon et al., 2014) as it can come with or without resulting symptoms (Dietz and Lanzarone, 2005, Otcenasek et al., 2007, Dietz et al., 2009). Levator ani muscle injury results in enlargement of the vaginal hiatus (Schwertner-Tiepelmann et al., 2012), reduced pelvic floor muscle strength (DeLancey et al., 2003,

DeLancey et al., 2007, Dietz and Shek, 2008b, Abdool et al., 2009a, Schwertner-Tiepelmann et al., 2012), or pelvic organ prolapse (DeLancey et al., 2007, Dietz and Simpson, 2008, Schwertner-Tiepelmann et al., 2012). Dietz et al. (2009) showed that levator avulsion is not associated with stress urinary incontinence, but Schwertner-Tiepelmann et al. (2012) showed a trend towards the development of faecal incontinence.

Longer term sequelae

The described pelvic floor symptoms after childbirth can be transient as can be seen from the diminishing symptom prevalences with time post partum in Iosif (1981), Viktrup et al. (1992) and Viktrup and Lose (2000). Likewise, physiological measurements can show pelvic floor restitution over time (Snooks et al., 1984, Cosner et al., 1991, Sultan et al., 1993, Tetzschner et al., 1996, Lee and Park, 2000). However, as Cornes et al. (1991) give to consider, we do not know whether this happens with natural restitution or whether women used any intervention.

Afshari et al. (2017) showed pelvic floor muscle strength in parous women within six months after delivery to be weaker than in nulliparous women. In women with at least one prior birth, pelvic floor muscle strength in the next pregnancy was statistically significantly lower when compared to primgravidae (Caroci et al., 2014). In Peschers et al. (1997) and Shek et al. (2012a), bladder neck descent and hiatal measurements did not show evidence of regression after two to three years post partum. A statistically significantly reduced pelvic floor muscle performance six to 11 years after vaginal delivery was demonstrated by Friedman et al. (2012), although some of the differences observed were small in magnitude.

Parous women have a higher prevalence of pelvic floor disorders later in life. Urinary incontinence for primiparous women has been shown to be associated with an adjusted odds ratio (OR) of around 1.3-1.6, for women with more than one delivery to be up to an adjusted OR of 1.5-2.0, although the numbers even out in old age (Milsom et al., 2017). Anal incontinence was associated with multiparity by an OR of 1.66 (95% CI [1.41–1.94]) in Matthews et al. (2013). For pelvic organ prolapse, the risk has been shown to be up to an OR of 5.3 in multiparous women (Rortveit et al., 2007, Milsom et al., 2017).

2.4 Pelvic floor muscle training in childbearing

If women experience enduring pelvic floor symptoms, midwives must refer them to the appropriate professionals for further management (International Confederation of Midwives,

2013). Depending on the diagnosis[25], the therapy will be determined. One first line rehabilitation[26] treatment for stress, urgency and mixed urinary incontinence in the postpartum[27] period is pelvic floor muscle training (Mørkved and Bø, 2014, Dumoulin et al., 2017, Woodley et al., 2017).[28] Training thereby means the systematic repetition of directed muscle contractions above threshold (Hollmann and Strüder, 2009); a pelvic floor muscle contraction refers to a concentric and isometric[29] voluntary inward and forward lift of the pelvic floor, squeezing around the urethra, vagina and anus (Messelink et al., 2005, Mørkved and Bø, 2015).[30]

The training must follow general muscle strength training principles of exercise science (Bø and Aschehoug, 2015). These are specificity (specific exercises for specific adaptation), progressive overload (gradual increase of effort that is higher than during daily activities), and maintenance. Thereby, dose response issues and adherence need to be considered (Alewijnse et al., 2007, Bø and Aschehoug, 2015). Enhancing muscle hypertrophy[31], muscle strength, power and endurance, enhancing stiffness of the connective tissue, and neuromuscular activation are goals associated with training (Fleck and Falkel, 1986, DiNubile, 1991, Ratamess, 2012). These morphologic and functional adaptations in muscle, nerve and connective tissues aim at increased muscle performance (Hollmann and Strüder, 2009, Ratamess, 2012).

The total workload (dose) of the training (type of exercises, frequency, intensity, duration and adherence) must be high enough (Mørkved and Bø, 2014). The current pelvic floor muscle training recommendation for urinary incontinence is 1-3 sets of 8-12 close-to-maximum contraction exercises per day, whereby a set is “a group of repetitions performed continuously without stopping or resting” (Bø, 2015b, p. 120). However, the term training in

[25] Apart from impairment of the pelvic floor muscles, injury to ligaments, nerves, or urethral structures may affect pelvic floor function (DeLancey, 2005, Dietz, 2013).
[26] Rehabilitation in this context is understood as the restoration of physical capabilities after a trauma (DocCheck Medical Services GmbH, 2018b).
[27] Pelvic floor muscle training as treatment in pregnancy is of uncertain benefit (Moore et al., 2013, Dumoulin et al., 2017, Woodley et al., 2017).
[28] For pelvic organ prolapse, anal incontinence, or to enhance sexual function, the intervention is of uncertain benefit during the childbearing time (Woodley et al., 2017, Wu et al., 2018).
[29] Concentric: muscle contraction reducing muscle length (Farlex, 2018c); isometric: muscle contraction maintaining constant muscle length (Farlex, 2018g).
[30] However, pelvic floor muscle work during increases in intraabdominal pressure in real life is eccentric, meaning that the muscle contracts while it is extended (Bø and Aschehoug, 2015, Farlex, 2018d). As eccentric muscle work can only be trained by eccentric exercises (Steiner, 2003) and eccentric exercises should be part of of any strength training programme (American College of Sports Medicine, 2009), such exercises might be important for pelvic floor muscle training, but no respective research is available (Bø and Aschehoug, 2015). Further, a trial to test involuntary reflexive pelvic floor muscle training in addition to this standard contraction training versus standard training alone for women with stress urinary incontinence is underway (Luginbuehl et al., 2015b).
[31] Increase in the size of muscle mass due to an increase in length and thickness of each muscle cell without an increase in the number of cells ('Muscular hypertrophy', 2007).

studies on pelvic floor muscle rehabilitation has also been used for training regimens that do not correspond to this contemporary recommendation; when reading about the topic, the meaning of the term needs to be clarified.

The theoretical objective behind recommending pelvic floor muscle training is that a stronger levator ani muscle better supports pelvic floor functions (Woodley et al., 2017). Training improves levator ani performance by enhancing muscle hypertrophy and stiffness (by increasing connective tissue mass), thus “elevating the levator plate to a permanent higher [anatomical] location inside the pelvis” (DiNubile, 1991, Bø, 2015b, p. 163). It also improves performance by enhancing neural factors to facilitate a more effective automatic neural response during increases in abdominal pressure and result in a faster contraction (Bø, 2004, Bø, 2015b) and perhaps urethral sphincter function (Moore et al., 2013). A second theoretical rationale for pelvic floor muscle training is that a conscious muscle contraction before and during increases in abdominal pressure, called the “Knack”, would avoid descent of the urethra and bladder base, increase urethral closure pressure and thus prevent leakage (Bø, 2004, Bø, 2015b). Dietz et al. (2009) however, who found no correlation between levator avulsion and urinary incontinence, raised the question whether pelvic floor muscle training might be effective for other reasons, e.g. training of surrounding muscular structures or unknown neuromuscular effects.

Pelvic floor muscle exercises seem to be a safe intervention, although the majority of studies did not consider adverse effects (National Collaborating Centre for Women's and Children's Health, 2013, Woodley et al., 2017). The only cited adverse effect in the systematic review by Woodley et al. (2017) is pelvic floor pain in two of 43 women in Stothers (2002). In Lagro-Janssen et al. (1992), pain (once) and an uncomfortable feeling during exercise (three times) were identified as adverse effects in 107 participants. Bø and Aschehoug (2015) reported slight headache, dizziness and discomfort for some women when starting with pelvic floor muscle training, probably due to an increase in blood pressure or inadequate breathing.

2.5 Preventive pelvic floor muscle training after childbirth

Pelvic floor muscle training is not only used for treatment purposes but also with the intention to improve pelvic floor muscle performance to *prevent* pelvic floor symptoms later in life. Intensive and supervised antepartum pelvic floor muscle training to prevent urinary incontinence up to three months post partum is a Grade A[32] recommendation for pregnant continent women (National Collaborating Centre for Women's and Children's Health, 2013,

[32] Grading of recommendations according to the Oxford system with Grade A as resting on best evidence being highest and Grade D being lowest (Centre for Evidence-Based Medicine, 2017).

Dumoulin et al., 2017). After childbirth, pelvic floor muscle exercises are routinely recommended to all women in postpartum healthcare practice (The Joanna Briggs Institute, 2011, Association of Chartered Physiotherapists in Women's Health, 2013a, Medizinische Kontinenzgesellschaft Österreich, 2016, Pelvic, Obstetric and Gynaecological Physiotherapy, 2017b). However, this usual care of oral or/and written recommendation of pelvic floor muscle exercises starting after birth to all women is not based on direct scientific evidence of effectiveness.

First, no study compared usual postpartum pelvic floor care, which differs between countries, with no exercising (Mørkved and Bø, 2014). Second, no studies have been performed on exclusively preventive pelvic floor muscle training starting after childbirth (Dumoulin et al., 2017, Woodley et al., 2017). Finally, preventive pelvic floor muscle training rests on the rationale that strengthening the pelvic floor soon after childbearing prevents incontinence later in life. Even if all women having given birth may be considered at risk for developing urinary incontinence later in life (Moore et al., 2013), a preventive benefit might not be the case as other factors seem to even out childbearing as aetiological factor for urinary incontinence in old age (Milsom et al., 2017). Mørkved and Bø (2015) argue that a long-term effect of pelvic floor muscle training is not to be expected as training needs to be continued to maintain the performance gain and to avoid detraining.

In view of the equivocal evidence, Dumoulin et al. (2017) and the NICE guidelines on postnatal care and urinary incontinence (Demott et al., 2006, National Collaborating Centre for Women's and Children's Health, 2013) do not give a recommendation for preventive pelvic floor muscle training starting after birth. The French College of Gynaecologists and Obstetricians (Senat et al., 2016), based on Deffieux et al. (2015), explicitly did neither recommend pelvic floor muscle training in the first two months after birth, nor for prevention in the medium and long term.

Looking at mixed prevention and treatment approaches including continent and incontinent women and starting after childbirth, 10 trials contributed to the analysis of self-reported urinary incontinence (Woodley et al., 2017). According to Woodley et al. (2017), there was, for the mid postpartum period up to six months, no difference (relative risk [RR] 0.95, 95% CI [0.75, 1.19]; 5 trials, 2800 women) in the prevalence of urinary incontinence (trials by Sleep and Grant, 1987, Meyer et al., 2001, Chiarelli and Cockburn, 2002, Hilde et al., 2013a, Kou et al., 2013). For the late postpartum period (more than six and up to 12 months after birth), there was “considerable uncertainty about the effect on urinary incontinence risk”, with a RR of 0.88, 95% CI [0.71, 1.09] (Woodley et al., 2017, p. 2); the very low-quality evidence stems

from three trials with 826 women (Meyer et al., 2001, Chiarelli and Cockburn, 2002, Kou et al., 2013).[33]

The findings of two of the three trials (Chiarelli and Cockburn, 2002, Kou et al., 2013) were in favour of pelvic floor muscle training. Chiarelli and Cockburn (2002) included women after vaginal operative delivery or with a baby ≥ 4000 g and applied adherence strategies. They found a statistically significant difference three months (OR 0.65, 95% CI [0.46, 0.91], p = .01) but not 12 months post partum (the most long-term outcome measurement taken (Chiarelli et al., 2004)). Both RCTs used a high intervention dose with professional supervision in the experimental intervention group (not reflecting usual care).

Only Meyer et al. (2001), Hilde et al. (2013a), and Kou et al. (2013) considered pelvic floor muscle performance by vaginal pressure and measured this by manometry, their results can be found in Table 1. According to Woodley et al. (2017), all three trials found no statistically significant difference for vaginal squeeze pressure between the groups; however, the cited values of Kou et al. (2013) show a statistically significant difference.

[33] Kou et al.'s (2013) data are published in Chinese language and only the English abstract could be read for this thesis; as this is not very informative, it is relied on the respective information cited by Woodley et al. (2017).

Table 1 Pelvic floor muscle performance in trials on mixed prevention and treatment of urinary incontinence post partum, complemented by effect on self-reported urinary incontinence

Study **Comparison**	**Outcome measure**	**Effect size pelvic floor muscle performance** Mean difference [95% CI]	**Effect size urinary incontinence** RR [95% CI]
Hilde et al. (2013a) Group comparison at postintervention test 6 months postpartum	Vaginal resting pressure (cm H_2O)	1.3 [–1.0, 3.6] *p* = .257	0.89 [0.60, 1.32] *p* = .569
	Vaginal squeeze pressure (strength) (cm H_2O)	3.3 [–1.4, 8.0] *p* = .172	
	Vaginal squeeze pressure (endurance) (cm H_2O seconds)	29.8 [–10.6, 70.2] *p* = .148	
Kou et al. (2013, cited in Woodley et al. (2017)) 12 months postpartum	Vaginal resting pressure (cm H_2O)	3.60 [–1.38, 8.58]	Mid-postnatal period 0.19 [0.04, 0.87] Late postnatal period 0.26 [0.08, 0.92]
	Vaginal squeeze pressure (strength) (cm H_2O)	26.10 [21.47, 30.73] Statistically significant	
	Vaginal squeeze pressure (endurance) (seconds)	1.80 [0.92, 2.68] Statistically significant	
Meyer et al. (2001) Group comparison at postintervention test 10 months postpartum	Vaginal squeeze pressure (strength) (cm H_2O)	–8.0 cm H_2O [–17.4, 1.4] *p* = .100 calculation by CO	Late postnatal period 0.82 [0.31, 2.21], calculation by Woodley et al. (2017)

Woodley et al. (2017, p. 44) came to the conclusion that "[t]he evidence to date about the benefit of mixed prevention and treatment approaches is […] not at all clear in postnatal populations." Dumoulin et al. (2017, p. 1461) concluded that "[h]ealth providers should carefully consider the cost/benefit of population based approaches to health professional taught […] postpartum [pelvic floor muscle training], that is, health professional instruction to all […] postpartum women regardless of their current or prior continence status" (Grade of recommendation: B).

There is also a variety of vaginal devices available for rehabilitation, serving as an external resistance against which muscles can contract. These devices can be used to perform pelvic floor muscle exercises around them and are e.g. anatomically formed (HOT Productions & Vertriebs GmbH, 2017), ball shaped with a long end that can be seen outside the vagina when the ball is inserted and whose movements can serve as an indicator of correct contractions (Jonasson et al., 1992, CampusPharma, no date), or with blades that have to be squeezed together (PT Direct, 2018). They may be weighted cones or balls to add weight onto the pelvic floor when inserted during being upright, with the intention to provoke involuntary or voluntary contractions to keep the device inside (Bø, 1995a, Herbison and Dean, 2013). Vaginal balls may additionally produce a vibrating effect (ELANEE, 2017b).

Summary

The pelvic floor as the bottom inner lining of the bony pelvis is made up of different muscles and connective tissue, with the main muscle being the levator ani. Amongst its various functions is to maintain continence, support the pelvic organs, and facilitate sexuality and childbirth, and the chapter presented the functional anatomy relevant for these tasks.

During pregnancy, the pelvic floor undergoes physiological changes to prepare the woman's body for childbirth, and vaginal childbirth is a most stressful event for the pelvic floor. In the postpartum time, the pelvic floor needs to rehabilitate from the challenges of pregnancy and birth. Dysfunction of the pelvic floor in pregnancy and after childbirth can lead to symptoms of urinary and anal incontinence, pelvic organ prolapse, and reduced sexual function. Anatomical and functional pelvic floor changes have also been shown for asymptomatic women. Symptoms or asymptomatic changes may be transient and subject to natural restitution, or enduring.

Pelvic floor muscle training is a therapeutic intervention for pelvic floor problems. It is also recommended in pregnancy and after childbirth with a preventive intention. However, the scientific evidence on a preventive effectiveness for continent women starting to train after childbirth is not firm. Various devices are available for rehabilitation purposes, for example vaginal cones or balls which are presented in more detail in the next chapter.

3 VAGINAL CONES OR BALLS FOR PELVIC FLOOR REHABILITATION, WITH A FOCUS ON THE POSTPARTUM PERIOD

In order to set the stage for the empirical part of this thesis, this chapter introduces the use of vaginal cones or balls for pelvic floor muscle rehabilitation. It first describes available devices and their methods of use, which is followed by theory about their working mechanisms. It then examines the scientific evidence around these devices and theoretical concerns about cone use. In order to clarify the knowledge base for performing a trial as the empirical work of this PhD, a systematic review was conducted on postpartum use of cones or balls. A summary of the systematic review is included in section 3.4 and closes this chapter.

Except for the systematic review, of which the search strategy is detailed in section 3.4.1, core textbooks, systematic reviews and articles served as first information sources on all topics (e.g. pelvic floor muscle training or performance measurement). Relevant and promising references therein were followed up and forward searches performed. The search for relevant vaginal devices and device information originated from the products known to the author at project start. Product information for these was looked up, and every product hint come across during the research was pursued.

3.1 Available devices and methods of use

A number of different cone or ball products is available, and the recommended methods for their use differ. These will be presented in the following two sections.

3.1.1 Nonvibrating cones or balls

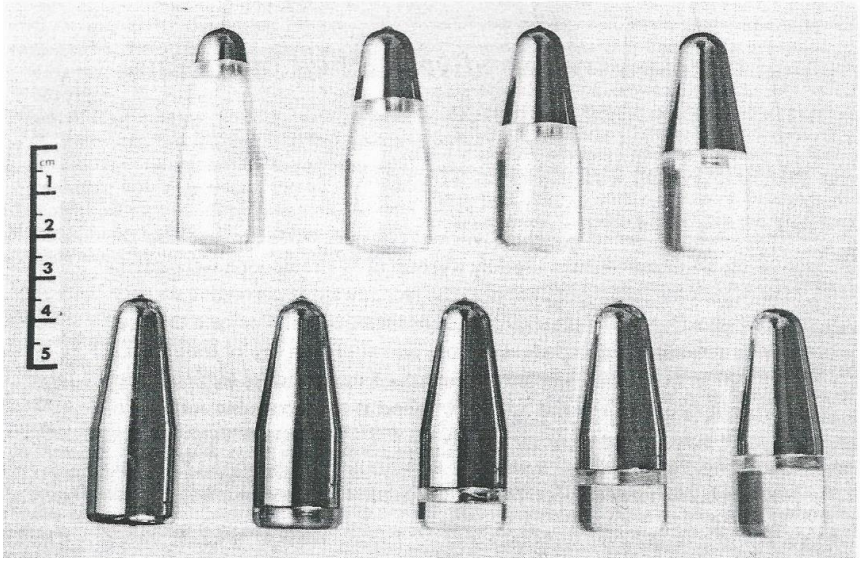
The use of cones to be inserted into the vagina for pelvic floor muscle exercising in modern science has been suggested by Plevnik (1985). Plevnik recommended a set of nine cones of equal volume but increasing weight from 20-100 g (Figure 8). The woman was advised to start with the heaviest cone that she could hold for a specified time, and to progress with the next heaviest cone when the so far heaviest cone became easy to hold.

Cones today come in different brands with different products, e.g.:

- ELANEE® Pelvic Floor Training Aids Phase I (ELANEE, 2017c): four cones from 20-71 g, advertised explicitly for the childbearing period,
- LadySystem® (Duchesnay, 2017): five cones from 5-55 g,

- Neen Aquaflex® (Patterson Medical, no date): one cone with exchangeable weights from 5-55 g, or
- VagaCare™ (Medgo, 1999): one cone with exchangeable weights from 30-90 g.

Figure 8 A set of vaginal cones as introduced by Plevnik



Source: Peattie et al. (1988).

According to Baumann (2012), cones were recommended in 60% of the German obstetric surgeries participating in her survey. Vaginal cones as helpful device after childbirth are mentioned in the professional midwifery literature (e.g. Varney et al., 2004, 'Femcon-Vaginalkonen', 2013)[34] and advertised on websites (e.g. ELANEE, 2017c, USMedia, 2017).

The use of vaginal cones was researched in RCTs either as a stand-alone intervention or in addition to pelvic floor muscle training, and compared to no treatment, pelvic floor muscle training (various regimens), or other therapies not of relevance here (Herbison and Dean, 2013, Dumoulin et al., 2017). In most studies, women were asked to insert the cone into the vagina and hold it for a specified time (10-45 minutes, once or twice per day or one to three times a week). This was done while moving around during daily life activities, or in a few studies during performing standardised activities such as coughing or lifting (Herbison and Dean, 2013). In this thesis, this method of cone use will be referred to as *hold use*. Within hold use, it can be differentiated between passive (no conscious effort) and active (contract

[34] It seems that the cones became less popular during the PhD period, as e.g. the cited advertisement in the Austrian Midwifery Journal has not appeared any more since 2013, and Royal College of Midwives website information about cone use available in 2013 is not available any more by now.

to hold) use (Peattie et al., 1988, Jonasson et al., 1989, Thakar and Stanton, 2000, Test-Club-Bericht, 2013). According to the muscle training principle of progressive overload, cone weight was usually increased during the treatment period so that the woman at the end of treatment was able to hold a heavier cone than at its beginning.

Some studies asked participants to contract around the cone during pelvic floor muscle exercises which in the following will be referred to as *contraction use*. A few studies researched hold use of cones with additional performance of pelvic floor muscle training (various regimens) without cone. Bø (2015b, p. 170) suggests that women could also "be asked to contract around the cone and simultaneously try to pull it out in lying or standing position, repeating this eight to 12 times in three series per day, or they can use the cones during progressively graded activities of daily living". Treatment duration ranged from four weeks to six months (Herbison and Dean, 2013, Dumoulin et al., 2017).

Although the term cones within contemporary urogynaecological literature mostly refers to cones as suggested by Plevnik (1985), it can, in contradiction to the original product and literal meaning of the word[35], also comprise similarly shaped devices or balls (Herbison and Dean, 2013, Bø and Aschehoug, 2015). According to Arvonen et al. (2002), ball shaped weights (e.g. Fishpond Ltd, 2018) might be more comfortable to use than cone shaped devices. In their study, the only one on nonvibrating weighted balls, Arvonen et al. (2001) researched both hold and contraction use of the device.

3.1.2 Vibrating balls

Balls (and one dumbbell shaped device) are also available in a vibrating form. Some of them are advertised explicitly for the childbearing period (ELANEE, 2017b, FUN FACTORY, no date-d, MAPA GmbH, no date-b), and they are discussed in online childbearing fora (e.g. Hebamme4U, no date, Powerfilm GbR, no date). They also appear in the professional literature (Glavind, 2001, Butej, 2010, Birk, 2012, Abdel Karim Ruiz et al., 2014, Heller, 2015, Porta-Roda et al., 2015, Rochera et al., 2017). When inserted into the vagina, a loose inner ball inside the device causes mechanical vibrations when the woman is moving around (not to be confounded with electrically produced vibrations). The balls are available in a single (see Figures 1 and 12 on pages 16 and 87) or double version, in different weights and from different manufacturers. Example products are:

- ELANEE Pelvic Floor Training Aid Phase II (ELANEE, 2017b): 44 g double ball consisting of a larger combined with a smaller ball with an inner ball in each,

[35] "[S]hape with a circular base and smooth curved sides ending in a point at the top" ('Definition von cone', 2019, https://www.collinsdictionary.com/de/worterbuch/englisch/cone).

- Fembowls (medesign I.C. GmbH, 2018a): 75-78 g double ball,
- Laselle™ Weighted Exercisers (Intimina, 2018d): three balls of same size (Figure 13 on page 87), weighing 28, 38, or 48 g; to be used as single or double (two combined single) balls,
- LELO Luna Beads™ (LELO, 2018): two balls of same size, weighing 28 and 37 g; to be used as single or double (two combined single) balls,
- NUK Pelvic Floor Trainer (MAPA GmbH, 2017, no date-b): dumbbell shaped (42 g),
- Smartballs (FUN FACTORY, no date-d, e, b): single ball Smartball uno (40 g, Figure 1 on page 16), double ball Smartball duo (72 g).

Of the named models, the Fembowls, the Laselle Weighted Exercisers and LELO Luna Beads are licensed medical devices as defined by the World Health Organization (2018b) (LELO, 2013b, a, ELANEE, 2017a, MAPA GmbH, 2017, FUN FACTORY, 2018, medesign I.C. GmbH, 2018b). Vibrating balls are sold in a healthcare context via pharmacies (e.g. Intimina, 2018a), chemists (e.g. Smartballs in Austria (BIPA, 2018)) or health retailers (e.g. LELO Luna Beads™ in the UK (Holland & Barrett Retail Limited, 2017)). They are also sold in an erotic context with the argument of enhancing sexual sensation by enhancing pelvic floor muscle performance (see e.g. Intimina (2018g, 2018b), 'Gib dir die Kugel!' (no date), medesign I.C. GmbH (no date-b), or FUN FACTORY (no date-c, no date-b)).

The recommendations for the use of vibrating balls vary. They can equal hold use of cones which aims to prevent the ball from slipping out and enhances ball weight during the weeks of use, or they can equal contraction use as described for nonvibrating weighted balls by Arvonen et al. (2001). As it is claimed that the vibrations would contribute to muscle strengthening (Schildbach, 2005, ELANEE, 2017b, FUN FACTORY, no date-d, MAPA GmbH, no date-b), there additionally can be the recommendation to use the vibration effect when moving around. In this thesis, this method of use will be referred to as *vibration use*. Table 2 gives an overview on recommended methods of use from different information sources.

The medical device licence and therefore potentially enhanced medical responsibility and vulnerability of the company might explain why the respective user informations of two of the three licensed medical devices instruct customers to perform exercises around the ball (Intimina, 2018c, LELO, no date); by recommending pelvic floor muscle exercises the companies stay on the safe side. However, although the (hold and) vibration effect mechanisms are more obviously advertised for those not licensed as medical device (ELANEE, 2017b, FUN FACTORY, no date-d, MAPA GmbH, no date-b), this is also the case for the medical device fembowls (medesign I.C. GmbH, no date-c), and hold and vibration

use is presented like an alternative to pelvic floor muscle exercises for the medical device Luna Beads (LELO, no date). A consumer journal article reporting on Laselle Weighted Exercisers did not discourage from hold use (Test-Club-Bericht, 2013). In this article, the vibrations are purported to regulate muscle tension and sensitise the vaginal wall, whereas the producer Intimina (2018d) informs that the vibrations help the woman to identify the correct placement of the ball.

Recommendations given in the professional literature are similar to those from the consumer information material and are also included in Table 2. In referring to Buchheit (1985), the German physiotherapist Heller (2015) mentions vibrating double balls called *ri-no-tama* balls. She recommends, among other methods of use, wearing the balls a few times daily (duration not specified) for postpartum pelvic floor muscle weakness with slight symptoms. She states that postnatal women are better able to accept the balls than the cones (which confirms Arvonen et al. (2002)), but does neither give a reason nor a source for this statement. Hold and vibration use of vibrating balls as sole form of therapy was researched by Glavind (2001) and was the intervention under scrutiny in Butej (2010). Porta-Roda et al. (2015) used vibrating balls during exercising the pelvic floor muscles (contraction use).

Table 2 Recommended methods of use of vibrating vaginal balls (from consumer literature to RCTs)

Source	Vibration use	Hold use	Enhance weight	Contraction use
Consumer information SMARTBALLS (no medical device) (FUN FACTORY, 2013, no date-d, c, a)	Yes	Yes	Yes	Not mentioned
Consumer information Luna Beads™ (licensed medical device) (LELO, 2013c, 2018, no date)	Yes	Yes	Yes	Yes
Consumer information Fembowls (licensed medical device) (medesign I.C. GmbH, no date-b, c, a)	Yes	Yes	Not mentioned (only double ball available)	Not mentioned
Consumer information Laselle™ Weighted Exercisers (licensed medical device) (Intimina, 2018b, c, no date)	Not mentioned (vibrations would allow the user to feel that the ball is correctly placed)	Yes in combination with contraction use	Yes	Yes
Consumer information ELANEE® Pelvic Floor Training Aid Phase II (no medical device) (ELANEE, 2017b)	Yes	Not mentioned	No (only one weight available)	Not mentioned
Consumer information NUK® Pelvic Floor Trainer (no medical device, dumbbell shaped) (MAPA GmbH, no date-b, c, a)	Yes	Yes	No (only one weight available)	Not mentioned
Butej (2010) (physiotherapy student final assignment; prevention of urinary incontinence, use of FUN FACTORY SMARTBALLS for 6 weeks)	Yes	Yes	No (uses double balls from the beginning)	Not mentioned
Heller (2015) (German teaching book on physiotherapy	Not mentioned	Yes (including Bø et al.'s	Not mentioned	Not mentioned

Source	Vibration use	Hold use	Enhance weight	Contraction use
post partum by German physiotherapist)		(2015b) version: draw on the retraction cord and at the same time try to keep the ball inside)		
Rochera et al. (2017) (article, use of “Geisha Balls” for physiotherapy)	Yes	Yes	Not mentioned but writes about one or two balls	Not mentioned
Abdel Karim Ruiz et al. (2014) (article about vibrating vaginal balls)	Yes	Yes	Yes	Yes
Glavind (2001) (scientific study, therapeutic use of “Geisha Balls” for urinary incontinence for 12 weeks)	Yes	Yes	Not mentioned (uses double balls from the beginning)	Not mentioned
Porta-Roda et al. (2015) Porta-Roda et al. (2013) (scientific study, therapy for urinary incontinence, use of pelvicGym™ balls for 6 months)	Not mentioned	Not mentioned	Not mentioned (uses double balls from the beginning)	Yes

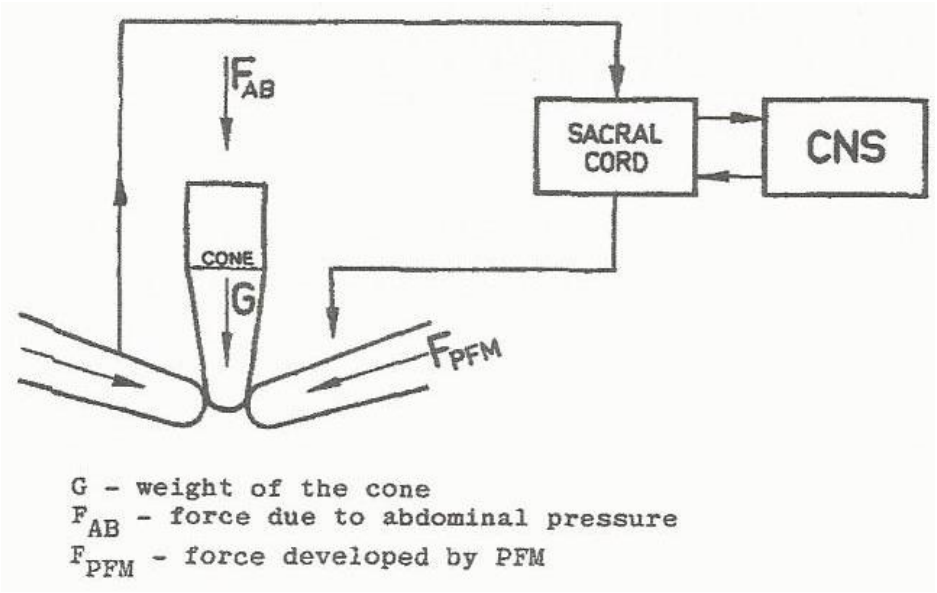
3.2 Theories about working mechanisms

The different methods of hold, contraction and vibration device use presented in the previous section have different assumed working mechanisms on how they would increase pelvic floor muscle peformance[36]. These are considered in the following and comprise physiological mechanisms and adherence.

3.2.1 Physiological mechanisms

Plevnik (1985) argued that the cone, by its weight, would tend to slip out of the vagina, and that this feeling of losing the cone would trigger a reflexive muscle contraction to retain it; increased intraabdominal pressure would enhance this reaction. Figure 9 explains his theoretical reasoning. Deindl et al. (1995) indeed demonstrated electromyographic[37] effects of cone use in the standing position, showing either an intermittent increasing and decreasing pelvic floor muscle activation pattern or activity with no variation.

Figure 9 Theoretical working mechanism of cones suggested by Plevnik (1985)



Source: Plevnik (1985)

[36] Apart from an influence on muscle performance, it is also purported that the balls would increase blood flow to the surrounding pelvic area (Schildbach, 2005, Abdel Karim Ruiz et al., 2014) and stimulate natural lubrication (Abdel Karim Ruiz et al., 2014). For a theory of vaginal acupressure by vibrating balls by Buchheit (1985), taken up by the company medesign I.C. GmbH (no date-c), see Appendix A.

[37] The recording of electrical activity in a muscle (Farlex, 2018e).

The underlying mechanism might be a contribution of the device to proprioception (Porta-Roda et al., 2015), the subconscious “internal sense of the relative position of the body’s musculoskeletal units with each other and the effort needed to move them” ('Proprioception', 2012, http://medical-dictionary.thefreedictionary.com/proprioception). This enhanced proprioception may then be used for biofeedback, a technique to teach an individual to consciously control automatic bodily functions with the help of equipment (Farlex, 2018b). Voluntarily holding back the device by squeezing the pelvic floor muscles around it provides biological signals that are returned (fed back) to the woman and indicate to her that her action has produced the desired physiological response. In this way, biofeedback teaches her to identify her pelvic floor muscles so that she is enabled to contract them consciously. The cones’ or balls’ biofeedback is sensory (by feeling pressure from the device) and kinaesthetic (by feeling the device move downwards) (Chiarelli and Moore, 2008). The named mechanisms, if effective, would support the basis of pelvic floor muscle training: awareness and voluntary control of the muscles (Baessler and Bell, 2008).

In contrast to Kegel’s (1948) concentric/isometric pelvic floor muscle exercises which would not adequately conform with exposure to everyday eccentric demands, reflex-related demands on the pelvic floor muscles by cones would be practised under and conform with everyday conditions (Fischer and Baessler, 1996, Luginbuehl et al., 2015b). Enhancing device weight over the time of use corresponds to the muscle training principle of overload (Ratamess, 2012). By performing pelvic floor muscle exercises around the inserted device or by drawing on the retraction cord and at the same time trying to keep the ball inside (as described on page 47), general strength training principles would be followed.

Compared to hold use only, Glavind (2001) suggested that the vibrations of vibrating balls would provide additional biofeedback. Rochera et al. (2017) claimed vibrioreceptors and baroreceptors provoking muscle contraction upon feeling vibrations, citing Abdel Karim Ruiz et al. (2014); Abdel Karim Ruiz et al. (2014) however did not write about receptors. When following up the idea, it is found that baroreceptors are to be found in blood vessels for the regulation of blood circulation (Shizgal and Hyman, 2013, DocCheck Medical Services GmbH, 2018a). Although mechanoreceptors able to feel vibrations can be found in tendinous sheets and in the urogeninal tissue (Schmidt and Schaible, 2006, DocCheck Medical Services GmbH, 2018c, Farlex, 2018i), these receptors are sensory only, not triggering muscle reactions. However, and this leads back to the above described propriception, proprioceptors as a kind of mechanoreceptors found in muscles and tendons sense movement, state of contraction and body position (Gardner and Johnson, 2013, Farlex, 2018k). Muscular strain reflexes triggered via these receptors’ signals (Silbernagl and

Despopoulos, 2012, Lehmann-Horn, 2017) might reflexively activate the pelvic floor muscles.[38]

Vibrations to improve muscle performance are used in whole body vibration therapy (Marín and Rhea, 2010) and in this context also have been of research interest for their potential effect on the pelvic floor (Lauper et al., 2009, Luginbuehl et al., 2012, Rodrigues et al., 2018). The vibrations are known to elicit isometric muscle contractions (in the thigh muscles) via reflexes and thereby increase muscle power (Rittweger et al., 2000, Bidonde et al., 2017). Therefore, this training theoretically could improve tone and strength of the pelvic floor muscles but was criticised for its lack of specificity to the pelvic floor (Baessler and Bell, 2008)[39]. This however can be provided by perineal applications (Rodrigues et al., 2018), and vaginal ball vibrations could also provide this specificity. However, whole body vibration training uses defined vibrations with frequencies of 12.5-30 Hertz and other specified technical parameters which cannot be achieved by a vibrating vaginal ball with its irregular mechanically induced vibrations; furthermore, the muscle strengthening effect is supposed to occur with higher frequencies (Baessler and Bell, 2008).

3.2.2 Adherence

In Porta-Roda et al.'s (2015) trial, contraction use of balls seemed to enhance adherence to pelvic floor muscle training in the long term, although this finding did not reach statistical significance. Contrary to this result, the urinary incontinence NICE guideline's conclusion was that the use of weighted vaginal cones is associated with more adherence problems if compared with pelvic floor muscle training (National Collaborating Centre for Women's and Children's Health, 2013). Herbison and Dean (2013) also pointed out that acceptability of cone treatment may not be optimal with a dropout rate of 22% of the 717 women included in their review who had used weighted cones (rate within included studies 0-72%).

All studies considered above randomised women to the interventions. It is possible that adherence by women who prefer such a method of pelvic floor muscle rehabilitation might be higher than that found in randomised groups, or higher than that to pelvic floor muscle training, and might thus enhance effectiveness. In Glavind's (2001) study, the (six) participants found it easy to use the balls–hold use saved time as the balls could be worn while performing other tasks.

[38] See Footnote 30 for an ongoing trial on reflexive pelvic floor muscle training in addition to contraction exercises.

[39] While in the two cited pelvic floor studies trainees were standing on a vibrating platform and this might be why Baessler and Bell (2008) claim a lack of specificity of whole body vibration therapy on the pelvic floor, the vibrations can also be applied more specifically to the pelvic floor muscles by the trainee sitting on the vibrating platform (Pearson, 2015).

3.3 Scientific evidence on device use (not in childbearing)

In order to gain an overview on the available scientific evidence on the use of vaginal cones or balls outside of the childbearing period, this section provides scientific information in terms of their effectiveness, limits of use, harms, and theoretical concerns.

3.3.1 Effect

The following sections present scientific evidence on nonvibrating and vibrating cone or ball use to treat or prevent stress urinary incontinence not (or not exclusively) in the postpartum period. Corresponding scientific evidence specifically for the postpartum period is discussed in the systematic review section 3.4.

Nonvibrating cones or balls

Herbison and Dean (2013) conducted a Cochrane review on weighted vaginal cones (including nonvibrating balls) for women with stress urinary incontinence. Twenty-three studies (only one of them–Wilson and Herbison (1998)–on the postpartum period) with 717 women were included, the protocol for cone use (cone weight, method of use, duration of use) differed across trials. The authors concluded that these devices seem to be better than no active treatment and of similar effectiveness to pelvic floor muscle training. However, they call the results of their review tentative and recommend larger, high-quality trials to reach a firmer conclusion on these devices' effectiveness. For the outcome pelvic floor muscle strength, no statistically significant difference was found between cones and controls (mean difference -1.19, 95% CI [-3.08, 0.71]) nor between cones and pelvic floor muscle training (mean difference -0.61, 95% CI [-2.49, 1.27]).

The NICE guideline on urinary incontinence (National Collaborating Centre for Women's and Children's Health, 2013) concluded (at evidence level[40] 1) that over the short term, vaginal cones are more effective than no treatment in women with stress urinary incontinence, and that there is no evidence of a difference in effectiveness between cones and pelvic floor muscle training (acknowledging that the appropriate training regimen for using vaginal cones is not clear). In Abrams et al. (2017), an international oeuvre on urinary incontinence, Dumoulin et al. (2017, p. 1498) conclude that "[vaginal cones] with supervised training sessions by a trained health professional may be offered as a first-line conservative therapy to those who can and are prepared to use them" (Grade B recommendation, see Footnote 32). In terms of efficiency, the National Collaborating Centre for Women's and Children's

[40] For intervention studies, the guideline defines levels of evidence from 1-4 (National Collaborating Centre for Women's and Children's Health, 2013). Level 1 (best) evidence stems from high-quality systematic reviews of RCTs or RCTs with a very low risk of bias, level 4 (weakest) evidence from expert opinion or formal consensus.

Health (2013) calculated that vaginal cones treatment is cheaper than pelvic floor muscle training, assuming the labour costs for using vaginal cones being one third of the costs for pelvic floor muscle training.

Salinas Casado et al. (1999) researched hold use of cones to strengthen the pelvic floor muscles in continent women in a study with a nonrandomised control group without treatment. The pelvic floor muscle strength measurement by digital palpation showed a statistically significant higher strength in the experimental group. However, the groups differed at baseline for some characteristics, and no statistical adjustment for these covariates was performed.

Vibrating balls

Two studies researched the use of vibrating vaginal balls. The first one was a trial with a single intervention group by Glavind (2001). The balls used were double balls, and their hold and vibration effect was studied. The duration of the intervention was 12 weeks with 15 minutes a day for one week and after that for half an hour a day while performing everyday tasks at home[41]. Of the 10 included women aged 34-59 years with stress urinary incontinence, six completed the study. Urinary incontinence measures were performed before treatment and after six and 12 weeks of treatment. Continence showed improvement, measured by a 24-hour pad test (measuring urine loss on a pad) and a questionnaire (not defined). Within its methodological limits (single group with few participants), the study showed encouraging results and the author recommended RCTs comparing vibrating balls with physiotherapy or with home exercises without physiotherapy.

Porta-Roda et al. (2015) researched contraction use of double balls by women with stress and stress-predominant mixed urinary incontinence over six months. As measured by the International Consultation on Incontinence Modular Questionnaire–Urinary Incontinence Short Form (ICIQ-UI SF) (Avery et al., 2004), there was an improvement for urinary incontinence which was more prompt in the ball group than in the control group. Measured by a one hour pad test, improvement was only shown in the ball group. No significant differences were found between groups via the King's Health Questionnaire (KHQ, Cardozo and Kelleher, 1997) and the subjective evaluation of efficacy. The authors concluded that the use of the balls together with Kegel exercises (contraction use) is effective in the treatment of mild to moderate stress urinary incontinence and that their use allows for more prompt positive results.

[41] As this could not be concluded with complete certainty from the report, Karen Glavind was contacted and confirmed that women did not have to perform any pelvic floor muscle exercises (Glavind, 2013).

3.3.2 Limits of use

Herbison and Dean (2013) reported that some women, e.g. those with a narrowed, scarred vagina, are physically unable to use cones. The NICE guideline on urinary incontinence formulated in more detail that vaginal cones are not suitable for all women: Their use is inappropriate when there is “a moderate to severe prolapse, too narrow or too capacious a vagina causing difficulty with insertion or misplacement of the cone, untreated atrophic vaginitis[42], vaginal infection, or during menstruation or pregnancy” (National Collaborating Centre for Women's and Children's Health, 2013, p. 84). From 22 trials, Dumoulin et al. (2017) summarised the reasons for attrition as low compliance, motivation problems, unpleasantness, aesthetic dislike, discomfort, and bleeding, whereby none of these reasons appeared predominant. All named limits assumedly also apply to vibrating balls. No further limits of use are specified in the studies about vibrating balls (Glavind, 2001, Porta-Roda et al., 2013).

3.3.3 Harms

To be able to judge the potential adverse effects of vibrating vaginal balls, all indications for adverse effects of vaginal cones or balls were followed up in the preparation of the present trial. In their systematic review on weighted vaginal cones, Herbison and Dean (2013) mention dislike of cones and bleeding as reasons for women dropping out of treatment. The NICE guideline on urinary incontinence (National Collaborating Centre for Women's and Children's Health, 2013, p. 84) cites Bø et al. (1999) who reported one woman with abdominal pain, one with bleeding and two with vaginitis (out of 29 using cones); another adverse effect identified with cones was occasional muscle soreness at initial use (Fischer et al., 1996).

In a German pharmaceutical consumer survey on the vibrating Laselle Weighted Exercisers (Test-Club-Bericht, 2013), no adverse effects were reported. However, this report needs to be regarded with caution as it was published in a consumer journal and might be biased for advertising purposes. Glavind (2001), using a model of vibrating vaginal double balls, described a slight vaginal irritation in one of her six (of 10) participants completing the study, and two of her participants found it unpleasant to put the balls into the vagina and discontinued participation. In Porta-Roda et al. (2015), mild local adverse effects were reported with ball use: hypersensitivity, irritation, itching and local discomfort. These disappeared as the study progressed, and all women continued participation; the authors concluded that the use of the balls together with Kegel exercises is safe. An Austrian

[42] Or vaginal atrophy: “thinning, drying and inflammation of the vaginal walls” due to low oestrogen levels (Mayo Foundation for Medical Education and Research, 2018, www.mayoclinic.org/diseases-conditions/vaginal-atrophy/home/ovc-20200167).

physiotherapist having been working with vibrating vaginal balls for 15 years confirmed that some women had experienced bleeding, but she never had seen a case of vaginal infection (Udier, 2014, see Footnote 10).

A closer look specifically at studies with postpartum populations showed that Wilson and Herbison (1998) did not report on adverse effects/harms other than two participants of unknown group withdrawing because they disliked treatment; nor did Norton and Baker (1990) and Jonasson et al. (1989) report any adverse effects.

3.3.4 Theoretical concerns

Bø (1995b), Bø (2015b) raised theoretical concerns about hold use of vaginal cones:

- It is not verified whether all women have the sensation of the cone slipping out as purported by Plevnik (1985). She cites Hahn et al. (1996), in whose study six of 30 incontinent women were able to retain the heaviest cone despite a weak pelvic floor. Closer examination by X-ray found the cones to be in a transverse position, indicating that these women did not feel the horizontally positioned cone slip out.

- This also suggests that cones may be held in place by other structures than the pelvic floor muscles. Hahn et al. (1996) explain the transverse lie with either a flaccid vaginal canal adapting to the weight of the cone, a vagina with a strong backward flexion enabling women to keep the cone in without assistance of the pelvic floor muscles, or by constipation with a filled rectum acting as a platform. It is unclear whether this argument also applies to ball shaped devices. Further, the influence of cysto- or rectocele on retaining the cone is not known.

- Cone use does not correspond to muscle strength training regimens or the exercise science perspective. Holding a cone for as long as 15–30 minutes may, as static or isometric muscle work, result in decreased blood supply and oxygen consumption, muscle fatigue and pain, and recruit other muscles instead of the pelvic floor muscles.

- Putting a weight on the pelvic floor muscles might fatigue instead of strengthen them; if putting weight above the pelvic floor muscles was effectively strengthening them, other factors leading to increased intraabdominal pressure (e.g. pregnancy) would also strengthen them. However, similar to Bø and Aschehoug's (2015) reasoning that a small increase in abdominal pressure may be an adequate stimulus for a pelvic floor co-contraction and thereby bring a training effect, a small cone weight could work as a training stimulus while a huge increase may weaken the pelvic floor muscles.

Despite her concerns, Bø (1995b) concluded that the effectiveness of cone therapy has to be evaluated in high quality studies. She also recommended more basic research to test whether most women have the sensation of the cone slipping out. Her theoretical concerns were taken up by Herbison and Dean (2013) who also concluded that further high-quality research on cones is needed.

The theoretical basis of this trial's research interest (hold and vibration use of vibrating vaginal balls) therefore cannot stand up against the theoretical reasoning of training science. However, to speak with Forman (1981), while theoretical understanding and explanation of phenomena is desirable and the goal of pure science research, the primary question in medicine is whether or not an intervention is effective. Herbison and Dean (2013) state that cones, albeit possibly not the best option, might still train the pelvic floor muscles and increase their strength.

3.4 Systematic review on the use of vaginal balls and cones post partum

Before embarking on a trial, a systematic review is recommended to inform the decision on whether it is necessary (Guyatt et al., 2008). A systematic review is a literature review for "a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review" (The Cochrane Collaboration, 2018, http://community.chochrane.org/glossary#letter-S).

For the present trial, this would have been, strictly seen, a quantitative systematic review on the effectiveness of hold and vibration use of vibrating vaginal balls to rehabilitate the pelvic floor muscles in the postpartum period. As from scoping searches it seemed likely that there would be no literature on this topic, and with the aim to produce a publishable systematic review, it was decided to widen the focus. The objective of the systematic review thus became to compare the effectiveness of vaginal balls or cones for improvement of pelvic floor muscle performance and urinary continence in the postpartum period to no treatment, placebo, sham treatment or active controls. A secondary objective was to gather information on effect on perineal descent or pelvic organ prolapse, adverse effects and economical aspects.

Up to 2013, no systematic review had focused on the use of these devices specifically during the postpartum period. Given the pelvic floor changes in the peripartum period, the progression in pelvic floor muscle performance improvement following rehabilitation measures during this time might differ from that outside of the childbearing time (Woodley et al., 2017), and the knowledge presented about vaginal cone use outside of the peripartum

period might not be directly applicable to women after childbirth. Accordingly, Demott et al. (2006) stated that the value of cone use in puerperal women is not clear.

The Cochrane review by Herbison and Dean (2013) had looked into the effectiveness of cones or balls for urinary incontinence and included one trial with postpartum participants. Another Cochrane review by Boyle et al. (2012) and a systematic review by Mørkved and Bø (2014) had looked into the effectiveness of pelvic floor muscle training during and after pregnancy. Boyle et al. (2012) did not include cones as intervention; Mørkved and Bø (2014) included cones amongst other forms of training but did not elaborate on this aspect in the results. Urinary incontinence was used as a primary outcome in all three reviews; studies having solely considered pelvic floor muscle strength as an outcome in continent women were excluded from the Cochrane reviews, whereas the use of this outcome was not made explicit in Mørkved and Bø (2014). Pelvic floor muscle strength in continent women as an outcome had been used in a systematic review on the prevention of pelvic floor dysfunction around childbirth by Harvey (2003). However, this review also only included cones amongst other pelvic floor muscle rehabilitation methods, it excluded studies on the treatment of urinary incontinence, and it was useful to search for more recent articles to update this review's findings.

Thus, a systematic review was needed which focused on (1) vaginal use of cones or balls as a pelvic floor muscle rehabilitation method (2) in the postpartum period, and (3) used both pelvic floor muscle performance and urinary (in)continence as primary outcomes to estimate effectiveness of device use. The systematic review was performed between October 2013 and October 2014 and published in *Midwifery* (Oblasser et al., 2015). This publication with all methodological details is reprinted in Appendix C; the following sections present an overview of the methodology and results of the systematic review performed to prepare the empirical part of this PhD.

3.4.1 Methodology

As required by The Cochrane Collaboration (Higgins and Green, 2011) and good practice for a research study, the systematic review was planned and outlined in a protocol. This was based on the guidance on systematic reviews of interventions by the Cochrane Collaboration (Higgins and Green, 2011) and registered at PROSPERO (international prospective register of systematic reviews in health and social care) (Oblasser et al., 2014a). During the systematic review, the protocol was slightly amended (because of feedback on the protocol

manuscript submitted for publication), and a version of the amended protocol was published in the *Journal of Advanced Nursing* (Oblasser et al., 2014b, reproduced in Appendix C)[43].

The objective of the review was to compare the effectiveness of vaginal balls or cones for improvement of pelvic floor muscle performance and urinary continence in the postpartum period to no treatment, placebo, sham treatment or active controls (e.g. pelvic floor muscle exercises). A secondary objective was to gather information on effect on perineal descent or pelvic organ prolapse, adverse effects, and economic aspects.

The research question was developed using the PICO (population – intervention – comparison – outcome)-framework outlined by O'Connor et al. (2011): *Does the vaginal use of cones or balls by women in the postpartum period improve performance of the pelvic floor muscles and urinary continence, compared to no treatment, placebo, sham treatment or active controls?*

Inclusion and exclusion criteria were developed on the basis of the PICOS (PICO plus study design)-scheme of the PRISMA Statement (Liberati et al., 2009). The types of study designs, participants, interventions, comparisons, and outcome measures, and report characteristics included in and excluded from this systematic review are listed in Box 1.

The search strategies used for each database are reproduced in Appendix D. Data were extracted from selected studies using a piloted standard data extraction form adapted from the data extraction form templates of the The Cochrane Pregnancy and Childbirth Group (no date) and The Cochrane Editorial Resources Committee (2013). The data abstraction form, including the scheme for quality appraisal, is enclosed as Appendix E. Before finishing this thesis, the search was repeated in pubmed to update the review's results. In this "2014-present" search, *sphere OR spheres* was added to the 2014 strategy as an additional search term for balls, and the restriction by the Medical Subject Heading (MeSH) *human* was dropped not to exclude the most recent and not yet indexed literature; the search also followed the link for "similar articles" to the published systematic review and feasibility trial protocol (Oblasser et al., 2016). Furthermore, the 2014 search strategy was repeated with the term *sphere OR spheres*. No new studies were found by the search update.

[43] After the manuscript had been accepted for publication, the amendments were registered at PROSPERO.

Box 1 Eligibility criteria for systematic review

Types of studies

Randomised and quasi-randomised controlled trials with individual or cluster randomisation and parallel design were included. Blinding of participants is not possible for this intervention.

Types of participants

- Women up to 1 year after childbirth at the time of beginning the intervention, of any parity, mode of birth and birth injuries, with or without urinary incontinence, were included.
- Pregnant women, women with anal incontinence or major genitourinary/pelvic morbidity were excluded.

Types of interventions

Vaginal use of cones or balls.

Inclusion criteria:

- cone or ball use of any frequency and duration, and of any method (combined with exercises or not)
- cones or balls of any form, size, weight or brand
- with any method of instruction (advised by any health practitioner or self-taught by information material).

Types of comparison

Comparison should be made with physiological restitution (no device or treatment) or any form of pelvic floor muscle training, e.g. physiotherapy individually or in group, or pelvic floor muscle exercises at home, with placebo or sham treatment.

Types of outcome measures

Outcomes should be measured immediately after the intervention, or be longer-term follow-up data.

<u>Primary outcomes</u>

Either one or both of these:

- pelvic floor muscle performance (e.g. strength, endurance), determined using a valid and reliable measure, e.g. vaginal squeeze pressure or participant reported improvement
- urinary (in)continence, determined using a valid and reliable measure, e.g. quantified symptoms or urodynamics.

<u>Secondary outcomes</u>

- perineal descent or pelvic organ prolapse as assessed by standardised clinical methods
- adverse effects, e.g. discomfort or pain during or after the intervention, or vaginitis, as determined in each of the included studies
- health economics, e.g. cost of interventions or teaching time, as determined in each of the included studies.

Report characteristics

There were no language, publication period or publication status restrictions.

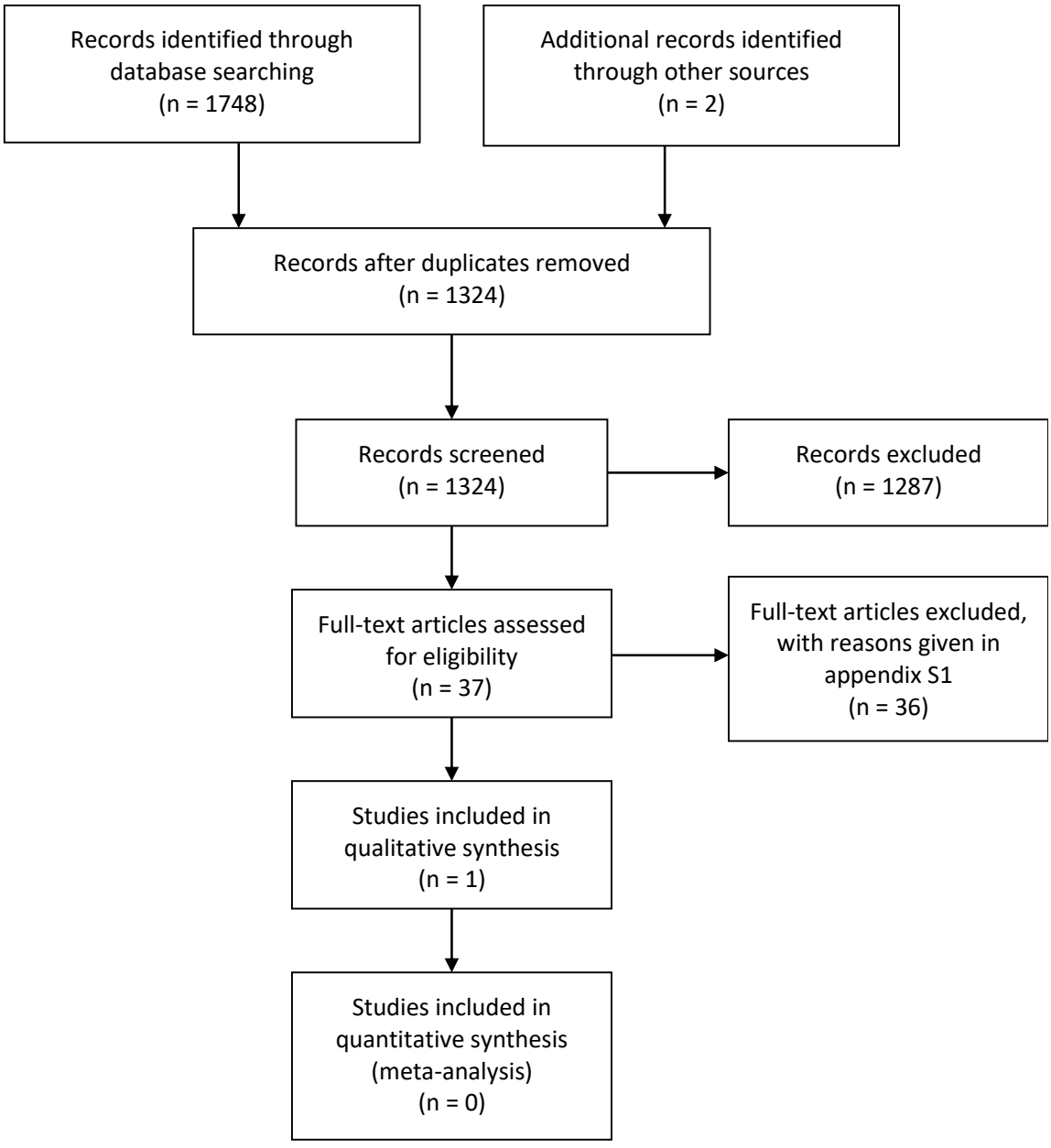
3.4.2 Result of systematic review

By the search technique used in 2014, 37 potentially useful articles were identified out of 1324 records screened. Vibrating devices had not been investigated for women after childbirth, and weighted cones had been researched in four studies (Jonasson et al., 1989, Norton and Baker, 1990, Fischer and Baessler, 1996, Fischer et al., 1996, Wilson and Herbison, 1998). Of these, only Wilson and Herbison's (1998) study was of sufficient scientific quality to be included in the systematic review.[44] As only one study was included, a data synthesis by meta-analysis was not possible and a narrative review was undertaken. However, a secondary intention-to-treat analysis on Wilson and Herbison's (1998) raw data enabled the researchers[45] to directly address the question of this systematic review in order to meet the systematic review objective. The PRISMA flow chart (Liberati et al., 2009) documents the literature assessment and selection process in Figure 10.

[44] The reasons for excluding the other three studies from the systematic review were poor methodology by nonrandom group allocation, use of an invalid and unreliable measurement method for pelvic floor muscle performance, and poor reporting. More detail on the excluded literature is given in Appendix F.

[45] CO as primary researcher, Janice Christie as second reviewer and supervisor, and Christine McCourt as supervisor.

Figure 10 PRISMA flow chart (according to Liberati et al. (2009))



Selected results of the reanalysis are shown in Table 3. Compared to the control group (routinely recommended pelvic floor muscle exercises), the cone group (hold use) shows a statistically significant lower rate of the primary outcome self-reported urinary incontinence at 12 months post partum (RR 0.63, p = .022), but an almost same rate of urinary incontinence in the cone group cannot be excluded (95% CI [0.40, 0.998]). Exploratory analyses of perineometry[46] measurements do not support the difference found for urinary incontinence (p values > .05 showing no statistically significant difference between cone and exercise group). Pelvic floor muscle strength results for the cone group compared to the control group show a wide mean difference 95% CI from -4.76 to 3.87 (point estimate -0.44), the muscle endurance mean difference 95% CI ranges from -2.11 to 4.39 (point estimate 1.14), indicating similar results and uncertainty.

Compared to the exercise group (enforced exercise regimen), the prevalence of urinary incontinence in the cone group is similar (RR 1.01, p = 1.000), but a prevalence of urinary incontinence half or almost twice as high in the cone group cannot be excluded (95% CI [0.52, 1.93]). Exploratory analyses of perineometry measurements support these findings (p values > .05). Pelvic floor muscle strength results for the cone group compared to the exercise group show a wide mean difference 95% CI from -6.87 to 5.01 (point estimate -0.93), the muscle endurance mean difference 95% CI results from -4.57 to 4.46 (point estimate -0.05), indicating similar results and uncertainty.

The trial had a high dropout rate, therefore it was important to consider the potential impact of dropout on the findings. This possible impact of dropout was recalculated by a sensitivity analysis as originally presented by Wilson and Herbison (1998). If all the participants who were not followed up were assumed to be incontinent, the prevalence of urinary incontinence would have been 81% in the control group, 69% in the cone group, and 74% in the exercise group. The group comparisons would then give the following results: cone group versus (vs) control group RR = 0.86 (95% CI [0.68, 1.08], χ^2 = 1.607, df = 1, p = .205), not showing any difference and effect of cone use; cone group vs exercise group RR = 0.93 (95% CI [0.70, 1.24], χ^2 = 0.047, df = 1, p = .829), not showing any difference between the treatments. If the participants who were not followed up were all assumed to be continent, the prevalence of urinary incontinence would have been 59% in the control group, 28% in the cone group, and 23% in the exercise group. The group comparisons would then give these results: cone group vs control group RR = 0.47 (95% CI [0.27, 0.81], χ^2 = 9.5, df = 1, p = .002), showing a greater effect of cone treatment than the complete case analysis; cone group vs exercise

[46] Vaginal manometric measurement of pelvic floor muscle pressure; for more information see section 5.4.2. Although the term perineometry is incorrect as no perineal pressure is measured (Peschers et al., 2001, Bø and Sherburn, 2005, Bø, 2015c), it is used in this thesis as common within the literature. It is used interchangeably with the alternative term vaginal manometry.

group RR = 1.20 (95% CI [0.55, 2.62], χ^2 = 0.041, *df* = 1, *p* = .840), not showing any difference between the treatments.

After 24-44 months and in women without further pregnancy or treatment, urinary incontinence shows a prevalence of 54% in the control group, 68% in the cone group, and 50% in the exercise group, but only 33% (in the respective groups 32%/53%/51%) of the original participants could be followed up. The cone group vs control group comparison gives a RR of 1.27 (95% CI [0.83, 1.94], χ^2 = 0.56, *df* = 1, *p* = .455), while the cone group vs exercise group comparison gives a RR of 1.37 (95% CI [0.80, 2.33], χ^2 = 0.71, *df* = 1, *p* = .399), not showing any differences between the groups. Pelvic floor muscle performance was not measured at this time point.

Overall, the review results suggest that hold use of cones could be helpful to treat urinary incontinence after childbirth. While an enhancing influence on pelvic floor muscle endurance seems possible, such an influence on pelvic floor muscle strength could not be shown. However, the validity of this intention-to-treat analysis is limited by the high rate of withdrawals, especially in the intervention groups, the lack of participant blinding, and by sample size not being based on a power calculation. Women after Caesarean section are included, and although the rate is similar in both groups, this could dilute the effect. Further, the fact that the experimental pelvic floor muscle exercise regimen in the trial does not correspond to contemporary exercise science recommendations risks to disadvantage the exercise group (the control regimen is not known), as do only four sessions with a physiotherapist from three to nine months after delivery as low intensity supervision compared to today's weekly supervision recommendation (Bø, 2015b). In contrast, the more intensive instruction in the cone and exercise groups might bias[47] the study towards a better result in these groups.

With respect to the effectiveness of vibrating vaginal balls, the review showed a dearth of scientific evidence as no study on the vaginal use of vibrating balls in the postpartum time could be identified. It was therefore concluded that further research is needed.[48]

[47] Bias is "systematic error, or deviation from the truth, in results or inferences" and "can lead to underestimation or overestimation of the true intervention effect" (Higgins and Altman, 2008, p. 188).
[48] In 2013, the question "Are vaginal cones an effective therapy for women with post natal stress urinary incontinence?" was contained within the no longer maintained UK Database of Uncertainties about the Effects of Treatments (UK DUETs; entry text available from CO).

Table 3 Results of systematic review reanalysis

Outcome	Cone group	Control group	Exercise group	Cone group vs control group (asks whether cones are better)	Cone group vs exercise group (asks whether groups are identical)
After 12 months					
	Prevalence			**RR [95% CI]**	
Self reported urinary incontinence	*n/N*=10/21 48%	*n/N*=69/91 76%	*n/N*=9/19 47%	0.63 [0.40, 0.998] *p* = .022 χ^2 = 5.25 *df* = 1	1.01 [0.52, 1.93] *p* = 1.000 χ^2 = 0.00 *df* = 1
	Mean (*SD*)			**MD [95% CI]**	
Pelvic floor muscle strength (perineometry, mean of three maximum strength contractions, cm H_2O)	*N*=19 12.7 (9.6)	*N*=79 13.1 (8.2)	*N*=19 13.6 (8.4)	-0.4 [-4.8, 3.9] *p* = .840 *t* = 0.20 *df* = 96	-0.9 [-6.9, 5.0] *p* = .750 *t* = 0.32 *df* = 36
Pelvic floor muscle endurance (perineometry, [mean of three?] contractions sustained over 5 seconds, cm H_2O)	*N*=19 7.8 (7.7)	*N*=79 6.7 (6.1)	*N*=19 7.9 (5.9)	1.1 [-2.1, 4.4] *p* = .490 *t* = -0.70 *df* = 96	-0.1 [-4.6, 4.5] *p* = .980 *t* = 0.02 *df* =36

Outcome	Cone group	Control group	Exercise group	Cone group vs control group (asks whether cones are better)	Cone group vs exercise group (asks whether groups are identical)
After 24-44 months					
	Prevalence			**RR [95% CI]**	
Urinary incontinence	*n/N*=13/19 68%	*n/N*=20/37 54%	*n/N*=10/20 50%	1.27 [0.83, 1.94] *p* = .460 χ^2 = 0.56 *df* = 1	1.37 [0.80, 2.33] *p* = .400 χ^2 = 0.71 *df* = 1

Note: *SD* = standard deviation; *df* = degrees of freedom.

Summary

One pelvic floor muscle rehabilitation method is the vaginal use of (vibrating) cones or balls. Various devices are available and used in different ways with different theoretical explanations. Their working mechanisms might lie in a better familiarisation with one's own pelvic floor by proprioception and biofeedback, in added resistance by device weight or during contractions, and in the vibrations; enhanced adherence to the intervention might also play a role. Although the intervention has its limitations and theoretical concerns about the use of vaginal weights for pelvic floor muscle rehabilitation exist, scientific evidence on their use outside of the childbearing period showed that the devices are potentially effective and safe.

A systematic review into the use of vaginal cones or balls for pelvic floor rehabilitation in the postpartum period revealed that the respective scientific evidence is very limited as only one study was found to fit the inclusion criteria. The results showed that hold use of cones may be helpful for urinary incontinence or to enhance pelvic floor muscle peformance after childbirth. No trial on the vaginal use of vibrating balls in the postpartum time could be identified. The next chapter thus provides the rationale for the ensuing empirical doctoral work and declares its research aims, questions and objectives, and the study design.

4 PURPOSE AND DESIGN OF EMPIRICAL THESIS PART

Concluding from the background knowledge about vaginal cones and balls elaborated in the previous chapter, and in order to clarify this thesis' purpose and design, this chapter first provides the rationale for this PhD's empirical research. It then goes on to describe, explain and justify the chosen research design. Finally, it declares the research aims, questions and objectives.

4.1 Rationale for the research

The research interest leading to this feasibility trial was to determine the effectiveness of hold and vibration use of vibrating vaginal balls to improve pelvic floor muscle performance in a postpartum population. The systematic review into the use of vaginal cones or balls for pelvic floor rehabilitation in the postpartum period and the wider literature review presented in Chapter 3 show that vibrating vaginal balls have only been tested in a single group of nonpostpartum women with urinary incontinence. The one RCT with a postpartum sample researched hold use of vaginal cones by women with urinary incontinence.

The thus identified knowledge gap calls for effectiveness research, whereby the best research design to determine the effectiveness of an intervention is an RCT (OCEBM Levels of Evidence Working Group, 2011). An RCT is justified when there is clinical equipoise, defined as a "situation in which it is not known which of two possibilities [X better than placebo; X worse than placebo] is more likely to be true" (Hulley et al., 2013, p. 333); this is the case for the intervention of interest.

The contemporary treatment development research approach, although focusing on drugs, asks for RCTs being performed before licensing and marketing treatments (NHS Choices Team, 2016, U.S. National Library of Medicine, 2017). As neither cones nor vibrating balls have been thoroughly researched for the specific postpartum group before being marketed, such research can be considered overdue from this point of view. Similarly, focusing on (physical) therapies, Bø and Herbert (2009) suggest a protocol for the implementation of new therapies into clinical practice which includes high-quality RCTs.

The novelty of the envisaged RCT lies in the researched (vibrating) device, in its method of use (holding it in the vagina during daily activities), in the targeted use by women after childbirth, the outcome pelvic floor muscle strength, and the study design. The aim of the planned future RCT to detect a potential difference in the change of the outcome values

between the experimental and a comparison intervention is expressed in the following research question:

> Does using vibrating vaginal balls differ from performing pelvic floor muscle training in its effect on pelvic floor muscle performance from pre- to postintervention measurements in women after childbirth?

To enhance rigour in a study, a pilot stage to test study elements is advised as part of the research process (Burns and Grove, 2007, Burns and Grove, 2009, Gerrish and Lathlean, 2015), and a good feasibility or pilot study improves the likelihood for successful trial completion (Hulley et al., 2013, National Institute for Health Research, 2017). Such preliminary studies are funded (NETSCC, 2016, ResearchNet, 2018, National Institutes of Health, no date), and funding streams might require a preliminary study to be presented to obtain funding for a full study (Craig et al., 2006).

Consequently, and with regard to the time and resources available, it was decided to perform such preparatory work for an intended RCT as the PhD project. Its research interest lies in assessing practical issues and feasibility of a future confirmatory RCT. It also integrates basic research on women's clinical experiences with the intervention as suggested by Bø (1995b). To understand the rationale for the chosen study design, the design issues will now be clarified.

4.2 Research design

The overall phrase to describe the design of this study is *single blind randomised controlled feasibility trial with two parallel groups*. The following sections will first clarify the RCT design and the nature of feasibility research. After this presentation of the methodological background to the present study, this particular study's design will be explained and justified in detail.

4.2.1 Randomised controlled trials

The RCT design is a form of experimental research design (Petrie and Sabin, 2009). It aims to determine a causal relationship between an intervention and an outcome (Hulley et al., 2013) and is the optimal design to answer the question: "Does this intervention help?" (OCEBM Levels of Evidence Working Group, 2011, www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf). The underlying paradigms of the RCT design are the paradigm of the experiment by Francis Bacon, the paradigm of repeated observation by David Hume, the paradigm of controlled comparison by John

Stewart Mill, and the paradigm of randomisation by Ronald Fisher (Kiene, 2001). Table 4 gives the rationale for each of these paradigms being used.

Table 4 RCT paradigms and their rationale according to the Institut für angewandte Erkenntnistheorie und medizinische Methodologie e.V. (2003)

Paradigm	Rationale
Experiment	Reliable empirical cognition is only possible in experimental conditions.
Repeated observation	A cause-effect relationship can never be determined on a single case but only by multiple repeated observations.
Controlled comparison	A comparison is needed to detect a causal relationship: the treated case against an untreated case (a multitude of treated cases compared to a multitude of untreated cases).
Randomisation	As the cases to compare mostly are not identical and react differently, the allocation to treatment or nontreatment must be done by randomisation.

Specific to an RCT therefore is the experimental comparison of a number of participants in at least two groups (Meinert, 2012). The classical RCT has a parallel group design (Hulley et al., 2013) where one of the groups serves as a control group, hence the term controlled. The allocation of participants to study groups is done at random (hence the term randomised) to avoid or reduce selection bias[49] (Higgins et al., 2011a) and to balance confounding[50] factors (Efird, 2011). To avoid or reduce performance[51] or detection[52] bias, the allocated intervention should be kept unknown to participants, carers and researchers involved in the study (Higgins et al., 2011a). This is called blinding or masking, and whereas the highest possible degree of blinding has to be strived for, it can be reduced to less than optimal levels for practical reasons (Meinert, 2012).

4.2.2 Feasibility research

The following sections provide the definition of feasibility research and inform about its design and methodological issues.

[49] Selection bias is a "systematic difference in characteristics between those who are selected for study and those who are not" (The Cochrane Collaboration, 2018, http://community.chochrane.org/glossary#letter-S).
[50] Confounding variables are associated with the outcome and the exposure (intervention) but are not an effect of the exposure (Jager et al., 2008).
[51] Systematic differences in exposure to other factors apart from the intervention of interets (The Cochrane Collaboration, 2018).
[52] Systematic differences in how outcomes are assessed (The Cochrane Collaboration, 2018).

Definition

A consensus definition of feasibility and pilot research in preparation for RCTs[53] was recently provided by Eldridge et al. (2016b). According to these authors, a "feasibility study asks whether something can be done, should we proceed with it, and if so, how. A pilot study asks the same questions but also has a specific design feature: in a pilot study a future study, or part of a future study, is conducted on a smaller scale" (Eldridge et al., 2016b, p. 1). They thus suggest the term feasibility as an overarching term for preliminary work with pilot studies being a subset of feasibility studies. Whereas pilot studies as (parts of) a planned main RCT on a smaller scale can be randomised or not, feasibility studies can have more varied approaches, such as a qualitative or survey design. Pilot studies can be external as a stand-alone study, or internal as the early stage of a definitive RCT.

However, for the inconsistent and interchangeable use of the terms pilot and feasibility in the scientific literature (Whitehead et al., 2014, Eldridge et al., 2016b), sources containing any of the terms were consulted for and are cited in this thesis. In citations, the terms used in the original work are kept.

Design and methodology issues

Feasibility research must show aims and objectives related to feasibility (Eldridge et al., 2016b). Shanyinde et al. (2011) looked at questions asked and answered in pilot and feasibility *randomised controlled* trials. Their summary of methodological issues that need evaluation in the context of an RCT is reproduced as Table 5.

A range of authors define further characteristics of pilot and feasibility studies (Thabane et al., 2010, Eldridge, 2013, Abbott, 2014, Whitehead et al., 2014, Lancaster, 2015, Eldridge et al., 2016b, National Institute for Health Research, no date-a, Williams and Lecouturier, no date). Amongst these are the need for prespecified feasibility criteria about study progression, a trajectory to the larger study and the use of the knowledge gained in the preliminary work to inform this future study, and appropriate thought about feasibility study sample size. To contribute to the development and evaluation of measures, measurement instruments and methods of data collection to be used in a full trial can be part of a pilot study's aims (Hulley et al., 2013, LoBiondo-Wood, 2013, Orsmond and Cohn, 2015, Indrayan, no date).

[53] Apart from testing the feasibility of a study, feasibility work can also test the feasibility of an intervention (Bowen et al., 2009, Shanyinde et al., 2011, Abbott, 2014).

Table 5 Methodological issues needing evaluation in the context of a pilot RCT (Shanyinde et al., 2011)

Issue	Needs to be evaluated in the context of a randomized pilot trial	Comments
Sample size calculation	✗	The numbers in a pilot RCT are unlikely to be adequate to get accurate estimates of effect size of variances.
Eligibility	✗	
Recruitment	✓	Referrals from clinicians are likely to depend on the RCT context.
Consent	✓	Consent rates in the RCT context are unlikely to be accurately estimated from asking about likely consent beforehand
Randomization procedures	✓	
Blinding procedures	✓	
Compliance/adherence to intervention	✗	Though, this could potentially depend on preference amongst interventions offered in the main trial
Acceptability of intervention	✗	Though, this could potentially depend on preference amongst interventions offered in the main trial
Cost and duration of intervention	✗	
Outcome assessment	✗	
Selection of most appropriate outcomes	✗	
Retention	✓	Retention may differ between experimental and control groups, and may depend on treatment preferences
Logistics of multi-centre trial	✓	
All components of the protocol work together	✓	

A feasibility study's analysis must follow its objectives. A descriptive analysis of the future trial's primary outcome measure can support the estimation of the sample size for the future RCT, which is a potential main aim of a feasibility study (Lancaster et al., 2004, Moore et al., 2011, Whitehead et al., 2014). Another use of descriptive analysis of potential primary outcome measures of the future trial can be to determine or design the most suitable outcome measure (Lancaster et al., 2004, Thabane et al., 2010, Shanyinde et al., 2011, Cocks and Torgerson, 2013, Williams and Lecouturier, no date).

The effect size calculation of the future trial's primary outcome measure is not considered a main objective of preliminary (feasibility/pilot) studies (Lancaster, 2015) and is recommended to be deemphasised (Arnold et al., 2009, Duan, 2013). Particularly, it is suggested to refrain from null hypothesis significance testing of the future trial's primary outcome measure with the argument that preliminary studies are not powered to draw firm conclusions in significance testing (Fidler, 2002, Lancaster et al., 2004, Leon et al., 2011, Eldridge and Kerry, 2012, Eldridge, 2013). Instead, a group comparison focusing on effect size and its CI can be performed to examine whether there is a promising (Stallard, 2012, Abbott, 2014, Ribeiro et al., 2014), potential (Arnold et al., 2009), or likely (Eldridge and Kerry, 2012) effect of the experimental intervention, and to see if further investigation of the intervention is worthwhile (Cocks and Torgerson, 2013, Abbott, 2014). Not the effect estimates themselves but the limits of their CIs shall be used in making any judgements in this matter (Eldridge and Kerry, 2012). It is recommended to interpret the calculated effect with caution (Lancaster et al., 2004, Arnold et al., 2009, Thabane et al., 2010, Duan, 2013, Abbott, 2014, Lancaster, 2015) "to avoid undue enthusiasm or pessimism about unstable estimates" (Arnold et al., 2009, p. S74).

4.2.3 A randomised controlled feasibility trial

In light of the described professional background on vibrating vaginal balls for pelvic floor rehabilitation in the postpartum period and the theoretical positions on RCTs and their preliminary research, the purpose and design features of the present study are presented in the following.

Research purpose

The purpose of this feasibility trial was to carefully prepare and test an intended RCT. The endpoint of this preparatory work was a trial design that is "fit for purpose" (Bugge et al., 2013, p. 9). Following on from the rationale for this research and from the information presented on the research design, the research questions of this study were:

- How can an RCT on the effectiveness of vibrating vaginal balls to improve pelvic floor muscle performance after childbirth be prepared and performed best?
- How many participants would be needed for a full trial to test effectiveness?
- Are there any harms associated with the experimental intervention?
- What are the participants' perspectives on and experiences with the interventions and the trial?

The study's aims and objectives were set according to recommended feasibility aims and objectives. The aims were as follows:

- Assess practical issues and feasibility of a future full RCT
- Monitor any possible harms of the experimental intervention
- Explore participant perspectives on and experiences with the interventions and the trial

The study objectives were as follows:

- Determine feasibility of a future full trial
- Explore different recruitment strategies
- Determine descriptive characteristics of the outcome measures
- Explore effect sizes
- Increase clinical experience with the experimental intervention
- Collect harms-related data of the experimental intervention
- Increase knowledge on the theoretical basis of the experimental intervention
- Investigate participant perspectives on and experiences with the interventions and the trial
- Determine adherence with the interventions and identify any adaptations needed to increase this.
- Assess staff, time, and budget necessary for a full RCT

Research design

The difference between a feasibility and a pilot trial is not clear cut. The National Institute for Health Research (no date-a, p. 1) defined a pilot trial as "a version of the main study that is run in miniature". Eldridge et al.'s (2016b, p. 1) definition of a pilot trial's specific design feature is that "in a pilot study a future study, or part of a future study, is conducted on a smaller scale". Although this trial was "the planned trial" (and not part of a planned trial) on a small scale, it was not the exact version of the planned full trial for the following reasons:

- All pelvic floor muscle performance outcomes were considered of equal importance as their testing was part of the feasibility trial. This means that no clinical outcome was designated as primary outcome as is recommended for a full RCT (Hulley et al., 2013).

- An aim of the trial was to test the scale used to measure participant reported pelvic floor muscle strength.

- The interviews served the feasibility trial purpose and are not applicable to a full trial.

- The online survey questionnaire was developed within this feasibility trial and tested for a full RCT.

- Using balls the same way as cones would necessitate initially choosing the heaviest ball the participant can hold and enhancing ball weight throughout the intervention period. Neither was done in this trial but might be done in a full trial in order to apply the experimental intervention in its potential fullest form.

Pondering the two cited definitions and the fact that the trial was not the exact version of a planned full trial, and considering further argumentation by Williams (2016), the decision was taken to label the present trial feasibility (and not pilot). As participants are allocated to the two parallel intervention groups at random, Arnold et al.'s (2009) suggestion to use the term *trial* was adopted.

Summary

The rationale for the present research was a lack of scientific evidence for the effectiveness of vibrating vaginal balls, a (medical) device sold and used in different countries, to strengthen the pelvic floor muscles after childbirth. An RCT is the optimal study design to evaluate intervention effectiveness, and a feasibility trial can prepare an RCT and examine its feasibility. The present single blind randomised controlled feasibility trial served to optimally prepare a future RCT, and its research questions, aims and objectives conformed to the feasibility study design. The next chapter considers the methods of this feasibility RCT.

5 METHODS

This chapter looks at each methodological component of this feasibility trial in order to explain how the trial was done. This starts with general feasibility trial methodology, goes on to sampling, recruitment and treatment assignment, continues with the interventions and outcomes with their measurement and data collection methods, and ends with data processing and analysis. Thereafter, ethical issues and PPI will be covered. For each topic, the choices made are described and justified.

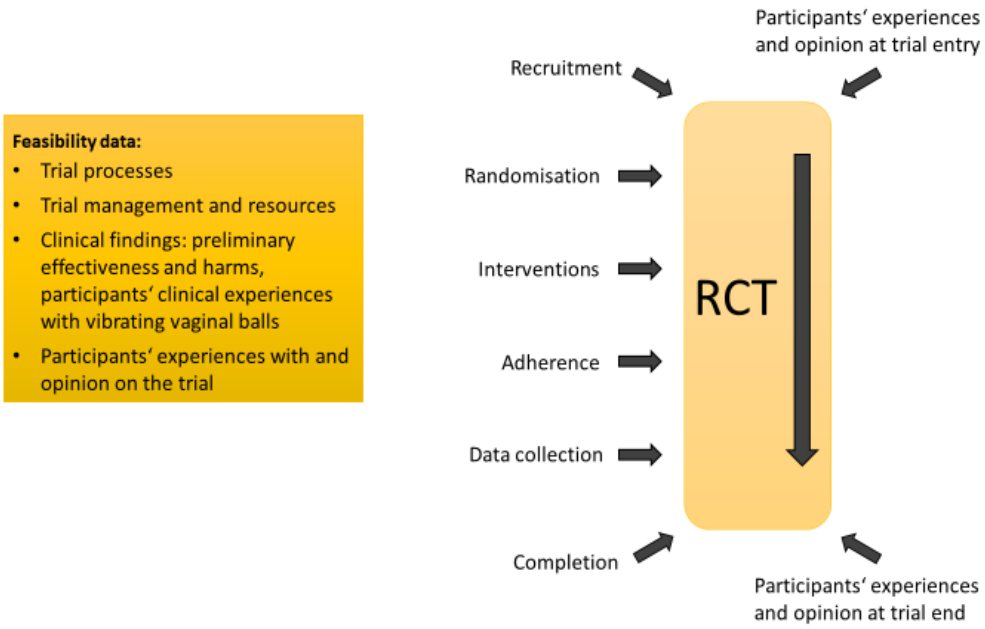
5.1 Feasibility trial methodology

Figure 11 shows the overall methodology of this feasibility RCT. Based on Thabane et al.'s (2010) framework, the feasibility of the future trial is determined for the trial processes, management and resources, and by the preliminary clinical[54] trial results. The framework is complemented by a survey on participants' experiences with and opinion on the interventions and the trial. This structure is reflected in the data collection and analysis sections of this methods chapter as well as in the results and discussion chapters. However, as per the feasibility study design, the trial level and the trial feasibility metalevel need to be considered for all issues, and as the feasibility is determined alongside the trial with the two levels being intertwined, the structure of this methods chapter follows that of the trial.

As methodologically required for a feasibility trial (Thabane et al., 2010, Abbott, 2014, Whitehead et al., 2014, Eldridge et al., 2016a), five feasibility criteria were set before trial start (Box 2). They were inspired by Cook et al. (2005) and represent educated estimates. It was determined in the research protocol that a full RCT would be deemed feasible if all criteria were met.

[54] Thabane et al. (2010) do not use the term *clinical* but *scientific*. However, the term *clinical*, defined as "involving or relating to the direct […] testing of patients" ('Definition of 'clinical'', 2018, https://www.collinsdictionary.com/dictionary/english/clinical), better encompasses the content and meaning of the respective outcomes in this feasibility trial.

Figure 11 Feasibility trial methodology of present study: an RCT is performed according to protocol and feasibility data collected alongside the RCT



Box 2 Feasibility criteria

Recruitment:

(1) At least 10% of eligible persons give consent to participate in the trial.

Preintervention data collection:

(2) At least 90% (95% CI [80, 100]) of participants attend the first pelvic floor muscle measurement within 3 weeks of consenting to take part in the trial.

Completion and adherence:

(3) At least 90% (95% CI [80, 100]) of participants start with the intervention within 4 days of the initial pelvic floor muscle measurement by perineometry.

(4) At least 80% (95% CI [70, 90]) of enrolled participants (a) keep to the assigned intervention group and (b) adhere adequately to the intervention.

Postintervention data collection:

(5) At least 80% (95% CI [70, 90]) of enrolled participants have the final data collection within 2 weeks of ending the intervention.

5.2 Sampling, recruitment and randomisation

After justifying the chosen sample size, the trial's inclusion and exclusion criteria and the sampling method will be explained. The different recruitment routes and the method of treatment allocation will be detailed.

5.2.1 Sample size

The sample size in feasibility studies should correspond to the feasibility objectives and be calculated a priori (Hooper, no date). As the feasibility rates are the primary outcomes in this feasibility trial, the sample size was calculated to be able to confirm these predetermined rates of 80% and 90% respectively with a margin of error of ± 10% with 95% confidence. Eldridge and Kerry (2012) and Hooper (no date) provide Formula 1 in Appendix G to calculate the width of the 95% CI for a single proportion given a sample size. This was converted to Equation 2 (Appendix G) for calculating sample sizes to determine proportions with a specified 95% CI. The decisive factor in this calculation is the lowest rate–in this feasibility trial 80% (for Feasibility criteria 4 and 5, Box 2). The determination of a rate of ≥ 80% with a 95% CI from 70-90 requires a sample size of 61. A sample size of 56 still gives a 95% CI of ± 10.48%.

Sim and Lewis (2012) also recommended a sample size of 55 for pilot studies to determine descriptive characteristics of interval and ratio outcome measures for an ensuing calculation of the sample size for a main trial. The gain of information is largest with a sample size of 55, with diminishing information gain from increasing sample size further.

Based on these two calculations, 56 women from the accessible population were recruited. Although other authors for various reasons suggested different numbers for feasibility/pilot trials (total sample size ranging from 20 to 150 or 3-9% of planned full RCT size (Browne, 1995, Julious, 2005, Hertzog, 2008, Stallard, 2012, Billingham et al., 2013, Cocks and Torgerson, 2013, Teare et al., 2014, Whitehead et al., 2016)), the chosen sample size was deemed appropriate as: 1) it was calculated to attain the feasibility proportions' 95% CIs with ± 10%; 2) Sim and Lewis' (2012) argument of no further information gain beyond a sample size of 55 seemed convincing; 3) it rather lies on the upper limit of the different suggestions; and 4) in addition to own data, perineometry data of other post partum studies are available for the consideration of outcome variability. It was also considered appropriate for pragmatic reasons, as the workload had to be handled within a PhD timeframe.

The sample size lies in line with the sample sizes of registered pilot and feasibility trials in the UK (median per arm in pilot trials 30 [range 8-114], in feasibility trials 36 [range 10-300] (Billingham et al., 2013)). Although inflation of sample size by 10-40% to account for loss to

follow-up is recommended (Thoma et al., 2010), it was not inflated in this project because of the feasibility design. Neither were the two lost participants replaced as it was calculated that 54 instead of 56 participants would only slightly widen the 80% feasibility rates' 95% CIs of ± 10% to ± 11% and thus be of little pragmatic significance.

5.2.2 Sample specification

This feasibility trial's eligibility criteria were determined with the intention to include healthy women experiencing physiological childbearing (according to the midwifery focus), minimise the risk of infection, control for confounding factors, and for pragmatic and ethical reasons. The inclusion and exclusion criteria with the rationale for their choice are given in Box 3.

Eligibility criteria checklists for recruiting professionals, CO's recruitment phone calls, and for the information/consent/initial study meeting can be found as Appendices H[55], I and J. Before trial entry, criteria were confirmed by consulting participants' maternity notes or by participants themselves, as specified in the information/consent/initial[56] study meeting schedule. For practical reasons, the ability to retain the ball in the vagina could only be determined after enrolment.

Although, as per the midwifery focus, a future RCT is intended to explore pelvic floor muscle strengthening to prevent (rather than treat) urinary incontinence, it was decided not to set urinary incontinence symptoms as an exclusion criterion in this feasibility trial for two reasons. The first reason was to gain an understanding of potential participants by finding out how many of the women wanting to participate in such a trial would have urinary incontinence symptoms. The second reason was the grey area of transient urinary incontinence after birth (Viktrup et al., 1992), when women with mild symptoms do not engage in therapy yet would profit from pelvic floor (self-)care. Women might have symptoms and wait for these to resolve on their own or with performing routinely recommended pelvic floor muscle exercises without professional surveillance which is the usual proceeding after childbirth in Austria. It was further thought that even symptomatic women might benefit from participation: By applying one of the interventions in the course of the trial, their pelvic floor muscles might be strengthened or pelvic floor symptoms reduced more than otherwise. Through the intensive focus on the pelvic floor, they might get a better pelvic floor perception and become more conscious about the importance of its health, which might contribute to their pelvic floor health in the long term. By raising the topic of incontinence or other symptoms, women with this problem might be identified and referred

[55] If applicable, Appendix documents are provided in German language (as used in trial) and as English translation which is denoted by the Appendix letter followed by (e), e.g. H(e).
[56] For clarity of expression, term has been changed from *first* meeting in research protocol to *initial* meeting in thesis.

(faster) to appropriate treatment who otherwise might be (more) reluctant to or only later (potentially when symptoms are enduring or worse) seek healthcare. It is known that women with urinary incontinence symptoms may wait many years before going to therapy and that worsening is the reason for their visit (Cammu et al., 2004).

Box 3 Eligibility criteria with rationale for their choice

Inclusion criteria
• Women from 6 weeks to 6 months after vaginal childbirth (at beginning of intervention) (I, C) • Term birth, i.e. 37+0 or more weeks of gestation (C) • Six weeks postpartum check by obstetrician performed and woman discharged from postpartum care with diagnostic findings appropriate to this period after childbirth (H, E) • Lochia have ceased (I) • Over the age of 18 with capacity to consent (E) • Sufficient knowledge of written and spoken German to be able to participate in the study (P, E) • Baby alive/not seriously ill (P)
Exclusion criteria
• Currently enrolled in pelvic floor muscle training with physiotherapist, midwife or fitness trainer (H, C) • Perineal tear 3rd or 4th degree at most recent birth (H, C) • Status post continence surgery (C) • Current pelvic floor or gynaecological surgery (H, C, I) • Current infection of genitourinary tract (H, I) • Recurrent (> 5 infectious episodes during last 12 months) or chronic (> 3 weeks duration of single episode in last 12 months) vaginal infections (C, I) • Neuromuscular conditions influencing pelvic floor muscle function (e.g. multiple sclerosis) (H, C) • Major medical condition influencing infectious risk (diabetes, immune suppressive therapy, HIV infection etc.) (H, I) • Currently on medication that could interfere with treatment or evaluation (C) • Currently enrolled in any other research study (P) • Pregnancy (also commencing during participation) or pregnancy planned within the study period (I, C) • Retention of ball is impossible (P) • Inability to perform the proposed procedures (P)

Note. Rationales: H = include healthy women, I = minimise the risk of infection, C = control for confounding, P = pragmatic, E = ethical (other than I).

5.2.3 Recruitment

Recruitment was performed via different routes and strategies (Figure 12). With reference to the researcher's workload, there was active recruitment where supportive maternity care professionals and CO approached eligible women, and passive recruitment where interested women could contact CO by themselves. The reason for this was to reach as many potential participants as possible, but also to explore recruitment routes and strategies in terms of success to identify suitable approaches for a future trial. "Recruitment success" thereby concerned the number and kind of women recruited, as it could be expected that different women would be reached via the different recruitment routes[57]. The general course of active and passive recruitment is described in the following sections, but as recruitment is also an outcome of this feasibility trial, detailed recruitment success and failure with regard to Feasibility criterion 1 is analysed and presented in the first process results chapter.

Active recruitment – contact initiation

Active recruitment was planned such that collaborating maternity care professionals would approach eligible clients with a brief introduction to the study and offer them a participant information and consent form[58] (Appendix K). If a woman was interested and accepted the information, she was asked for her verbal consent to have her name and telephone number passed on to CO. After communication of the completed recruitment form from the professional, CO called the woman to introduce herself, ask her if she was interested in participation, and to answer any queries she had. If a woman's interest and eligibility was confirmed at this initial contact, a date either for the information/consent/initial study meeting or for another phone call (if a woman wanted more time to reflect or if there was any reason to postpone) was arranged. The active recruitment process was documented on the documentation sheet for active recruitment contacts (Appendix L).

Active recruitment paths led via the delivery suite of the AKH Vienna and community midwives and obstetricians' surgeries in Vienna; planned recruitment via the postnatal wards of the AKH Vienna did not succeed. Process details on recruitment via these different routes are provided in Appendix M.

Passive recruitment – contact initiation

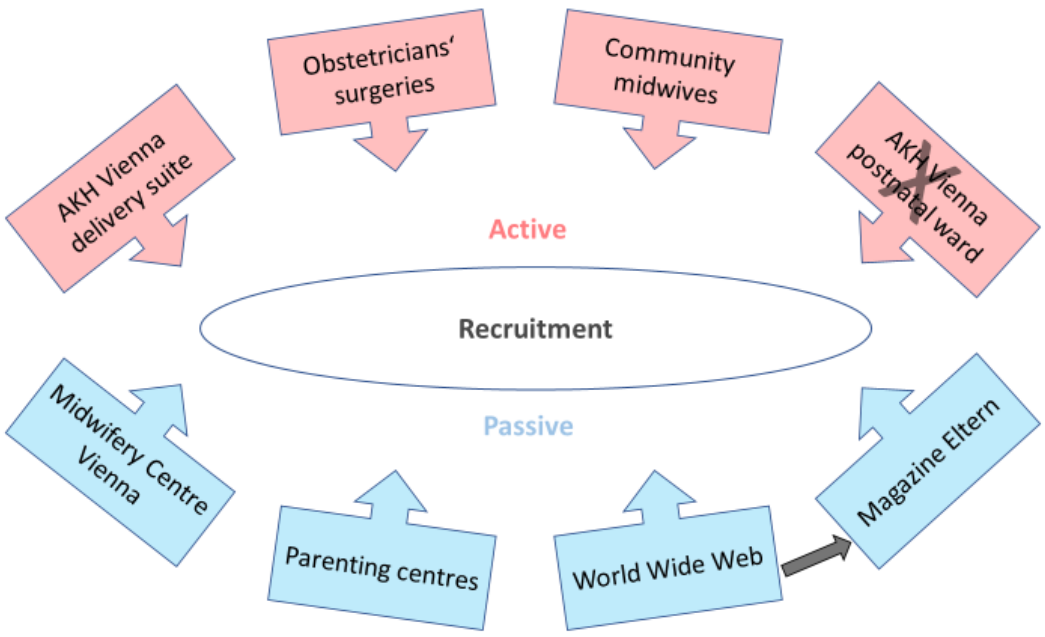
Passive recruitment means that women had somewhere found the trial recruitment text (see recruitment sheet in Appendix 13 a/b of the research protocol [Appendix N]) or heard or read

[57] Explanation for UK readers: The Austrian health care system is not as centralised as the UK health care system, hence the various recruitment routes.
[58] As required by the Ethics Committee of the Medical University of Vienna, the participant information sheet was combined with the consent form.

about the trial and contacted the researcher on their own initiative. If a woman's basic eligibility (e.g. vaginal birth at term, not attending a postpartum exercise class) and further interest was confirmed at this initial contact with CO, she was sent the participant information and consent form via e-mail. To allow time to reflect on participation, the potential participant was called not earlier than 7-10 days after she had been sent the form to ask whether she was still interested, to answer possible questions, check eligibility and to set a date for the information/consent/initial study meeting if applicable.

Passive recruitment sites were the Midwifery Centre Vienna (Hebammenzentrum Wien), a private parenting centre (Nanaya) and a parenting centre of the City of Vienna, the World Wide Web, and the parenting magazine *Eltern*. Process details on recruitment via these different routes are provided in Appendix M. The passive recruitment process was documented on the documentation sheet for passive recruitment contacts (Appendix L).

Figure 12 Recruitment paths



Recruitment via the AKH Vienna postnatal ward did not work, and via the parenting magazine *Eltern* develped out of recruitment at the World Wide Web.

Recruitment in both routes following contact initiation

After the described phone contact between CO and potential participants, recruitment continued in the same way for all recruitment paths. If a woman was interested in participation and eligible, a personal meeting between her and CO at the venue of her choice (either at her home or at a public place) was arranged. At this meeting, the feasibility trial was fully explained (including visual material) and details were clarified. If applicable, the participant information and consent form was then signed. After having assigned a participant identification number (ID), data collection started.

5.2.4 Sampling method and sample labelling

In active recruitment, the most easy to reach potential participants were approached. This constitutes opportunity sampling (Burns, 2000)[59]. In passive recruitment, participants selected themselves by contacting CO on their own initiative; the resulting sample is thus named a self-selection or volunteer sample (Lund Research Ltd, 2012) which can also be seen as a form of opportunity sampling (Burns, 2000, Trochim, 2006b, LoBiondo-Wood and Haber, 2013a). The different recruitment venues should result in a hospital- and community-based sample.

5.2.5 Treatment assignment and blinding

The randomisation was done in randomly created blocks of different size. Blocked (also termed restricted (Schulz and Grimes, 2002)) randomisation contributes to a more even distribution of participants between groups than simple randomisation and is especially important in small samples; randomly created blocks of different size avoid selection bias by keeping the investigator blinded to the size of each block and thus unable to predict allocation (Efird, 2011). To obtain more information about the experimental intervention, the allocation was skewed with two thirds of participants (n = 37) being allocated to the experimental and one third (n = 19) to the comparison arm (according to Eldridge and Kerry, 2012, Duan, 2013). As no significance testing was performed in this trial, the unequal group size is not relevant.

The randomisation sequence was produced via Sealed envelope™ (Sealed Envelope Ltd, 2018), an online randomisation database for clinical trials. To avoid selection bias, correct randomisation conceals the allocation sequence from the persons involved in the study before and when it is done (Higgins et al., 2011a); the list therefore was created by a fellow PhD student and accessible only to her until disclosure to CO after completed randomisation. Calculation instructions fed into the programme were: 56 participants;

[59] Alternatively termed convenience, accidental, or haphazard sample/sampling (Trochim, 2006b).

allocation to Group A, Group A, Group B for a 2:1 allocation ratio; and block sizes of 3, 6 and 9. This prompted the programme to produce a table with 63 participants (with the same information entered, created table size varied from 57 to 63 depending on the created blocks). As only 56 participants were needed, the randomisation list was cut after number 56. This resulted in the planned allocation of 37 participants to the experimental group and 19 participants to the comparison group.

Opaque sealed envelopes serially numbered with the participant IDs and containing randomisation slips with the codes generated from the computer programme were prepared by the PhD colleague for CO. Each time a participant was to be randomised, the envelope with her ID was opened by her or in her presence. Blinding of participants (by withholding information about the alternative option) would have been difficult to achieve and made the trial flow more complicated with the given interventions, and there would have been a large risk of disclosure as participants met and talked to each other at the measurement appointments; thus, participants were not blinded. CO, who was allocating participants to groups, explaining procedures, collecting baseline and participant reported outcome (PRO) data and conducting interviews, was obviously not blinded for logistical reasons. However, the trial is called single blind because the perineometric pelvic floor muscle measurement was done by blinded external assessors. With this procedure, the highest possible standard of blinding was maintained.

5.3 Interventions

This feasibility trial had two intervention groups, with participants either using a vibrating vaginal ball as the experimental intervention (but not discouraging them from their standard pelvic floor muscle exercises) or performing pelvic floor muscle training as the comparison intervention. Although it would have been an option to ask the experimental group to do both interventions, or to include a third group doing both interventions (as in Wilson and Herbison (1998)), the decision for these two groups only was taken for two reasons. The first was the interest to research vaginal ball use not as an addendum but a possible alternative to pelvic floor muscle exercises (like in the majority of cone trials (Herbison and Dean, 2013)), the second to ease study performance with regard to this being a feasibility trial. It was not an option to include a group without any intervention as recommending pelvic floor muscle exercises constitutes standard care after birth which for ethical reasons must not be withdrawn from study participants (Meinert, 2012).

To avoid bias, it was ensured that apart from necessary specific differences according to allocated intervention (e.g. questions about the experience with the intervention), the groups

were kept identical in terms of study performance and data collection. The interventions in detail are described and justified in the following.

Experimental intervention

The experimental intervention details corresponded to previous trial protocols, which were amended for pragmatic and logistical reasons. Participants were encouraged to use the vibrating vaginal ball Laselle Weighted Exerciser (at trial start named Laselle Kegel Exerciser) 28 g from Intimina/LELO (Intimina, 2018d, Figure 12). The ball was inserted (if desired with some lubricant) with the lower end about 2 cm deep into the vagina to lie just above the levator ani muscle. It was left for 15 minutes once daily during the first week to see if it was well tolerated (meaning that no difficulties were reported by participant at first adverse event call). This was then followed by wearing the ball for 30 minutes once daily from the second week onwards. With the intention to achieve the vibrating effect, the ball was worn while moving–e.g. performing everyday tasks or going for a walk. Detailed instructions for use were given to participants in the information and consent form (Appendix K) and explained verbally by the researcher at the information/consent/initial study meeting.

Figure 13 Vibrating vaginal ball used: Laselle Weighted Exerciser 28 g

Source: Intimina (2018e). Item reprinted with permission of Intimina.

The ball should not be worn during menstruation and intercourse; all other contraindications were covered within the exclusion criteria (section 5.2.2). They were formulated according to various ball and cone information materials and experiences from studies on ball and cone use.

The particular product Laselle Weighted Exerciser was chosen for three reasons. First, the author was advised to use a medical device in line with the Austrian law on medical devices (Bundeskanzleramt Rechtsinformationssystem, 1996) and the Laselle Weighted Exercisers fulfil this requirement (license provided in Appendix 29 of research protocol [Appendix N]). Second, they were the choice of most PPI participants when shown different products and they were, third, suggested (as used) by an Austrian physiotherapist.

Although Laselle Weighted Exercisers weighing 38 and 48 g are available and any two balls can be tied together to result in combined weight, only the 28 g ball was chosen for this feasibility trial and ball weight not enhanced during the intervention period. This is in contrast to most cone trials (Herbison and Dean, 2013), but was, like the number of groups, decided with the intention to keep the feasibility trial design simple. Finding the heaviest ball a participant could hold or enhancing ball weight would also have incurred costs for additional balls and surpassed the available budget. Nevertheless, a small weight may also challenge the pelvic floor muscles and at the same time forego the theoretical concern about too heavy a weight being to the disadvantage of the muscles (Bø, 1995b, Bø, 2015b).

The daily duration and frequency of ball use was set in correspondence with popular use in practice which is between 30 minutes and some hours once daily (Butej, 2010, FUN FACTORY, 2013, Rochera et al., 2017, 'Gib dir die Kugel!', no date, FUN FACTORY, no date-d, f, Intimina, no date, medesign I.C. GmbH, no date-c). In two studies, cones were also worn for 30 and 45 minutes respectively (Arvonen et al., 2002, Castro et al., 2008), although in many cone studies the cones were used 15 minutes twice daily (Herbison and Dean, 2013). However, whereas the cone trials used cones as a therapeutic intervention for urinary incontinence and cone application therefore might need to be more considerate of deficient muscles, this feasibility trial aimed at women with no pelvic floor symptoms. One application daily was considered easier than two in terms of practicability for women with little children.

As it is ethically required to use a study treatment in addition to standard care if some form of standard care is available (Meinert, 2012), participants in the experimental group were not discouraged from performing pelvic floor muscle exercises on their own as this is standard care in Austria. However, they were not encouraged to perform the exercises either as the trial did not aim to research the use of balls in addition to pelvic floor muscle exercises. Participants were informed orally about this at the information/consent/initial study meeting and asked at the final interview whether they had been performing exercises.

Comparison intervention

The underlying question of interest in this research is whether using vibrating vaginal balls is more effective than natural restoration for pelvic floor muscle strengthening after childbirth. However, this comparison is not amenable to research because it is not ethical to withdraw an existing standard treatment (Meinert, 2012) which in this case is recommending pelvic floor muscle exercises. Therefore, the comparison group was advised to follow standard care.

Standard care after birth in Austria is the routine recommendation of pelvic floor muscle exercises. It was planned to encourage participants in this group to continue or start the pelvic floor muscle exercises they had been recommended by customary written instructions from their health professionals after birth. In case participants would not have been given written instruction by their caregivers and as an Austrian standard on pelvic floor muscle training after childbirth was not available at the time of trial planning[60], an instruction sheet (Appendix 2 in participant information and consent form [Appendix K]) was prepared according to the pelvic floor care guidelines after childbirth by the UK Association of Chartered Physiotherapists in Women's Health (2013b, a)[61]. As it turned out during the trial, any written information participants had been provided with was tailored to the immediate postpartum time[62]. This therefore led to all participants in the comparison group getting the instruction prepared for this trial, as they all were at least 6 weeks post partum. It asked participants to perform three blocks of pelvic floor muscle exercises daily in different body positions and with each block containing 8-10 long and 8-10 short contractions. The detailed instruction sheet was given to participants in the information and consent form (Appendix K) and explained by CO at the information/consent/initial study meeting.

There are no contraindications given for pelvic floor muscle exercises in the extended postpartum period, only the recommendation to adapt them to the individual pelvic floor condition (Boyle et al., 2012, Heller, 2015). They seem to be a safe intervention, detailed safety information is provided in Appendix 8 of the research protocol (Appendix N).

Both groups

Duration of intervention use was set for 12 weeks to match the recommended pelvic floor muscle training duration in the NICE guideline for urinary incontinence (National

[60] There still is no standard in 2018, but in 2016, a pan-Austrian information sheet on continence in childbearing (including exercise instructions) was created by the Medical Continence Society Austria (Medizinische Kontinenzgesellschaft Österreich, 2016).
[61] Now named Pelvic, Obstetric and Gynaecological Physiotherapy (POGP)
[62] Such as e.g. the leaflet by the German physiotherapist Angela Heller (no date) distributed at the AKH Vienna.

Collaborating Centre for Women's and Children's Health, 2013). These 12 weeks are endorsed by a physiological perspective: In the first eight weeks, training mainly entails neural adaptations, whereas the desired muscle hypertrophy needs more time to develop (Saltin, 1986, cited in Bø et al., 1990a, DiNubile, 1991). Twelve weeks are also in accordance with Culligan et al. (2010) who aimed at strengthening the pelvic floor muscles in women with little or no pelvic floor dysfunction, and with Glavind's (2001) vibrating vaginal ball study. In cone studies, intervention duration was between 4 weeks and 6 months with a mode of also 3 months (Herbison and Dean, 2013).

All participants were explained the "Knack", a conscious contraction before and during increases in abdominal pressure to support pelvic floor function (see section 2.4). Intervention diaries used in both groups to chart adherence (see next section) were also intended to enhance participants' adherence with their intervention (as in Sleep and Grant, 1987, and Mason et al., 2001a) and possibly retention within the study, as were the phone calls for adverse events monitoring (as in Peattie and Plevnik, 1988, and Gorbea Chavez et al., 2004). These phone calls also served to answer queries.

5.4 Data collection

This section first introduces the collected data in an overview and a graphical representation of the flow of data collection from a participant's view. This is followed by details on the data collection process for all data collected.

5.4.1 Overview

The data collection comprised feasibility measures as primary outcomes, and clinical measurements and women's perspectives and experiences as secondary outcomes. Table 6 gives an overview on the data collected together with their method of measurement and measurement timepoints.

The data collection period ran from February 2015 to May 2016. All data were collected by CO, except vaginal manometry data which were collected by blinded assessors, and adherence which was charted by the participants themselves. For the different data collected, 13 different data collection forms were used (some of them in two versions for the two intervention groups), and Table 6 refers to where the respective forms can be found in the appendices.

Table 6 Overview of data collection

Data	As measured by	Timepoint	Data collection form
Primary Outcomes			
Feasibility	Recruitment rate	Within 4 weeks of ending recruitment	n.a. as calculated from information on different forms
	Timely preintervention pelvic floor muscle measurement attendance rate		
	Timely start of intervention rate		
	Retention rate		
	Timely postintervention data collection rate		
	Adherence rate		
	Adherence chart	Daily at time of intervention (12 weeks)	Appendices O and P
	Trial process observation/description	At time of ongoing trial data collection	Informal
	Trial management observation/description		
	Resources necessary		Appendix 12 of research protocol (Appendix N)

Data	As measured by	Timepoint	Data collection form
Participant characteristics			
Demographic and clinical baseline variables	Structured interview	At time of initial study meeting, before randomisation	Appendix 19 of research protocol (Appendix N)
Potential mediating or moderating variables	Structured interview	At time of final study meeting	Questions in final interview Appendices Q and R
Secondary outcomes			
Participant reported pelvic floor outcomes	Structured pelvic floor questionnaire	Within 3 weeks before the intervention	Appendices 20a/b of research protocol (Appendix N)
		Within 2 weeks after the intervention	Appendices 20a/b of research protocol (Appendix N)
Externally assessed pelvic floor muscle performance	Vaginal manometry measurement	Within 4 days before the intervention	Appendix S
		Within 2 weeks after the intervention	
Type, severity and number of adverse events	Active and passive surveillance (phone and end interview, self-report)	At time of intervention (12 weeks)	Appendix T
		Within 2 weeks after the intervention	Questions in final interview Appendices Q and R
	Active surveillance (final call)	Within 2 weeks after final data collection	Appendix T
	Passive surveillance (self-report)	Up to 1 year after end of participation	n.a.

Data	As measured by	Timepoint	Data collection form
Participant perspectives on and experiences with interventions and trial	Structured interview long form – for first 18 participants of experimental and first 10 participants of comparison group (together 28)	Within 3 weeks before the intervention	Appendix U
	Structured interview long form – same 28 participants as for previous outcome measure	Within 2 weeks after the intervention	Appendix Q
	Structured interview short form – the 28 of 56 not included in the two previous outcome measures	Within 3 weeks before the intervention	Appendix V
	Structured interview short form – same 28 participants as for previous outcome measure	Within 2 weeks after the intervention	Appendix R
	Structured anonymous online survey questionnaire – same 28 participants as for previous outcome measure	Within 1 day after final data collection with reminders 2 and 4 weeks later (where applicable)	Appendices W and X

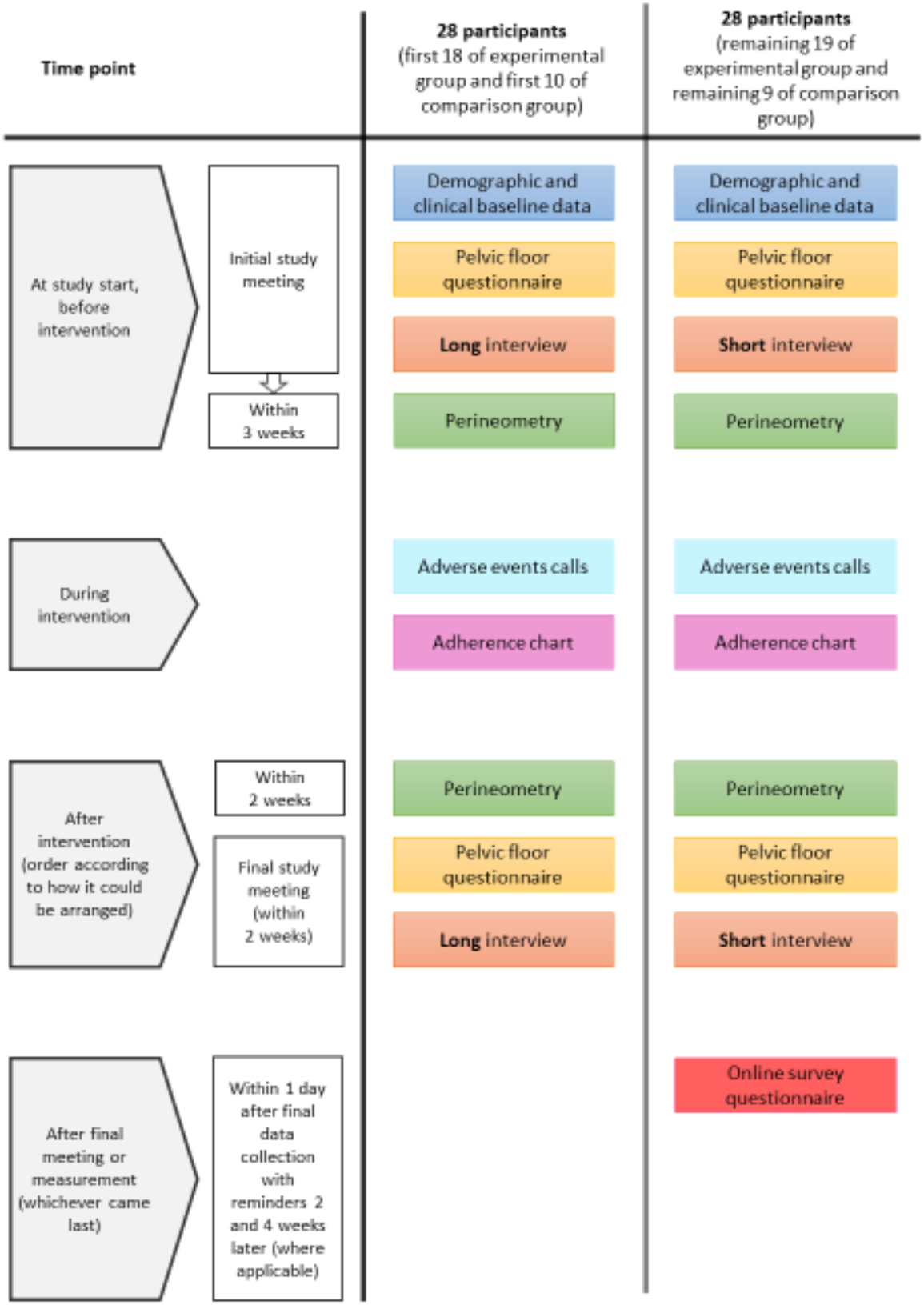
Note. Envisaged timepoints and participant numbers according to research protocol.

Figure 14 visualises the data collection with the focus on the participants. The illustration is inspired by Perera et al.'s (2007) recommendation on reporting complex interventions, adapted here to report a complex data collection process. The same information (valid at the trial planning stage) can be found in the format of the SPIRIT guidelines (Chan et al., 2013) in the published research protocol (Appendix C).

The information/consent/initial study meeting was a personal meeting of CO with each potential participant, on the venue most suitable to the woman (her home or public place where sufficient privacy could be obtained) and of approximately 1.5 hours duration. The final study meeting was a personal meeting of CO with each participant, again on the venue most suitable to her, with a duration of approximately 45 minutes. Usually, no other person besides the participant (and her children) and CO was present in the room. Appendix J shows the content and flow of the information/consent/initial study meeting, Appendix Y the final meeting checklist.

For the technical pelvic floor measurement by a blinded assessor with an approximate duration of 15 minutes, participants had to travel to the AKH Vienna where the assessors work(ed) or studied and where a room and equipment were available to perform the measurements. All participants got a final phone call to thank them and say good bye. Detailed information on data collection is laid out in the following sections.

Figure 14 Flow of data collection from participants' point of view



Note: Envisaged time points and participant numbers according to research protocol.

5.4.2 Details of data collection

As summarised above, the data collection comprised trial process, management and resources outcomes, participants' demographic and clinical baseline data, clinical outcomes and a survey on participant experiences and opinion. For each of these areas, collection details are given under the following headings, clarifying the kind of data collected, the justification for their collection, and the techniques used to collect them.

As measurement instruments must be valid[63] and reliable (LoBiondo-Wood and Haber, 2013b), validity, reliability and other important measurement features are considered (following the specifications by Hulley et al. (2013)). Respective information available at the trial planning stage is provided. Information gained on the measurements' functionality, appropriateness and feasibility during the course of the trial is reported in the results sections.

Trial process data

Trial process data cover the feasibility assessment of the key processes that will be part of a full RCT (Thabane et al., 2010, Tickle-Degnen, 2013). This includes the areas of participant selection, sampling and recruitment, randomisation, the interventions with adherence, and data collection.

Assessment of these processes was done in quantitative terms by the feasibility criteria. To be able to calculate the specified feasibility rates, relevant details on recruitment, pre- and postintervention data collection, trial completion and adherence were documented. Feasibility assessment of trial processes in qualitative terms was informed by collecting the reasons why women declined participation, by informal observation, by describing and reflecting on the processes from the researcher's point of view, and by extracting relevant information from the participants' perspectives on and experiences with the trial. The assessors commented on the measurements in the electronic measurement form and at a final meeting.

In the case of pelvic floor muscle training, adherence is a factor influencing effect, and it is known that adherence to pelvic floor muscle training is a challenge (Alewijnse et al., 2007). When adherence could influence trial effect, it should be charted and taken into account as a mediating[64] factor (Goetghebeur and Shapiro, 1996). Different ways to define and measure adequate adherence were used in pelvic floor muscle training studies (Mason et al., 2001a,

[63] Stengel (2010) also name responsiveness (change sensitivity) but Hays and Hadorn (1992) demonstrate, in theoretical terms, responsiveness to be an aspect of validity.
[64] A mediator is a variable which "represents the generative mechanism through which the focal independent variable is able to influence the dependent variable of interest" (Baron and Kenny, 1986, p. 1173).

Mørkved and Bø, 2015, Woodley et al., 2017). Studies having measured adherence to pelvic floor muscle training during childbearing (as outlined by Woodley et al. (2017), Mørkved and Bø (2014) and Mørkved and Bø (2015)) were screened for their method used to do so. Except in Bø et al. (1999), no method of adherence measurement was reported in cone studies (Peattie et al., 1988, Seo et al., 2004, Herbison and Dean, 2013).

From the identified options, it was decided to use an intervention diary in which participants were asked to chart their adherence (Appendices O and P). As defined in the research protocol (Appendix N), adherence was quantified as adequate if the prescribed intervention sessions had been completed in at least 80% of the days of the intervention period, and as inadequate below that (as in Braekken et al., 2010); an intervention session was completed if either the prescribed intervention dose or the personal training maximum until fatigue was reached. For participants in the experimental group, the days of menstruation were subtracted from the preset 80% threshold of 84 days as participants were instructed to not use the ball during menstruation. For each participant to whom that applied, her individual 80% threshold was set (in days) and adherence calculated from there. As menstruation is no contraindication to pelvic floor muscle exercises, this did not influence adherence calculation in the comparison group. Participants were also interviewed about the reasons keeping them from being adherent.

When taking a closer look at the charts of nonadherent participants, it was found that 14 of those in the experimental group had noted slightly less than 15 and 30 minutes of ball use respectively without marking it as maximum performance, e.g. 28 instead of 30 minutes. To not unduly give weight to minor protocol deviations, margins of error as shown in Table 7 were set for an adherence sensitivity analysis. In the comparison group, a number of participants had performed two daily exercise blocks (of two sets each) instead of three. According to Bø (2015b) and the NICE guideline (National Collaborating Centre for Women's and Children's Health, 2013), the current evidence-based recommendations on pelvic floor muscle training for urinary incontinence are one to three sets of 8-12 contraction exercises per day and at least three sets of eight contractions respectively. To not unduly give weight to potentially insignificant protocol deviations, margins of error as shown in Table 7 were also set for this intervention group for an adherence sensitivity analysis.

Table 7 Margins of error set for adherence sensitivity analysis with ensuing gain in participants labelled as adherent

Experimental group	Comparison group
Small margin of error: > 10 instead of 15 minutes ≥ 25 instead of 30 minutes ➔ plus two adherent participants Wider margin of error: ≥ 10 instead of 15 minutes ≥ 20 instead of 30 minutes ➔ plus four adherent participants	Performing exercise blocks two times/day instead of three times/day ➔ plus three adherent participants

To further determine the interventions' feasibility, participants' experiences with the interventions were collected in the interviews at study end, and the respective methodology is described in the section on participants' experiences and opinion on page 108.

The trial process assessment data collection was self-designed without formal validity and reliability evaluation available. Nevertheless, it was considered appropriate, sensitive, and specific, but with room for subjective perception and interpretation. Construct, content and face validity[65] seemed to be given.

Trial management data

Collecting and analysing management data should serve to identify respective problems to optimise management of a full trial (Thabane et al., 2010). Management data comprise human, organisational and data management issues (Simon, 2010, Thabane et al., 2010, Tickle-Degnen, 2013), in this feasibility trial e.g. the organisation of the rota for the blinded pelvic floor assessors, the logistics of the devices used, or ethical considerations. The trial management assessment data collection was informal, following the suggested questions of the named authors and of Orsmond and Cohn (2015). Issues related to management arising during the trial, including those brought up by the pelvic floor assessors and participants, were noted. Although there is room for subjective interpretation, the method used seemed appropriate, sensitive, and specific, and to have content and face validity.

Resources data

To be able to assess the resources necessary for implementing a full RCT, details on staff, facilities and material needed (Thabane et al., 2010, Tickle-Degnen, 2013) were collected

[65] Construct validiy "is the degree to which a specific measuring device agrees with a theoretical construct"; face validity "describes whether the measurement seems inherently reasonable"; content validity "examines how well the measurement represents all aspects of the phenomena under study" (Hulley et al., 2013, p. 39).

during this feasibility trial, together with the collected materials' receipts. The form in Appendix 12 of the research protocol (Appendix N) to document staff tasks with time needed and material items with quantity helped with this. This resources assessment seemed appropriate although the form of data collection was self-designed without formal evaluation available.

Participant characteristics

Demographic and clinical baseline data shall describe a sample at enrolment and enable detection of respective differences between randomised groups to identify potential confounders (Meinert, 2012). Therefore, participants' demographic and clinical characteristics, including a number of potential confounders to pelvic floor health (e.g. ethnicity, body mass index [BMI], smoking, parity, mode of delivery, birth weight, or previous birth injury (Milsom et al., 2017)), were collected for each participant before randomisation. This was done by structured interview with the participant and/or by checking the maternity notes where relevant. The full range of collected data can be seen in the respective data collection form in Appendix 19 of the research protocol (Appendix N).

Motivation was rated highly by physiotherapists as good prognostic feature to improve stress urinary incontinence with physiotherapy (Mantle and Versi, 1991), and motivation and adherence appear to be associated with positive outcome (Bø and Larsen, 1992, Dumoulin et al., 2014, Bø, 2015b). As, according to Cammu et al. (2004), there is no validated measure for motivation, it was measured in this feasibility trial via a numerical rating scale from 0-10 (McDowell, 2006, Harpe, 2015), with rating being asked during the initial interview. This should find out participants' motivation to adhere to the interventions during the trial and to compare the motivation between the intervention groups.

In the final interview, information on two potential mediating variables was collected. One was postnatal exercise/pelvic floor class attendance during the trial, the other whether participants in the experimental group had performed pelvic floor muscle exercises. Further, the moderator[66] breastfeeding status was collected. The method used for collection of participant characteristics seemed appropriate, sensitive, and specific, and to have content and face validity.

Clinical outcome data

The collection of clinical outcome data aimed to gather participants' clinical experiences with vibrating vaginal balls. It should also determine the feasibility of the collection of the future trial's clinical outcomes, estimate their descriptives, and assess preliminary intervention

[66] A moderator is a "variable that affects the direction and/or strength of the relation between an indpendent or predictor variable and a dependent or criterion variable" (Baron and Kenny, 1986, p. 1174).

effect sizes and harms (Thabane et al., 2010, Tickle-Degnen, 2013). This supports researchers in deciding whether it is justified to progress to a full trial (Stallard, 2012) and contributes to an informed calculation of full trial sample size (Lancaster et al., 2004, Thabane et al., 2010, Moore et al., 2011, Eldridge, 2013, Tickle-Degnen, 2013, National Institute for Health Research, no date-a, Williams and Lecouturier, no date). The collection details of the clinical data are presented in the following.

Participants' clinical experience with vibrating vaginal balls

To gain more understanding on vibrating vaginal ball use and its influence on the pelvic floor (as suggeted by Bø (1995b)), participants' clinical experience with the experimental intervention was researched with four questions on pelvic floor reactions. These data were collected in the final interviews, and the respective methodology is described in the section on participants' experiences and opinion on page 108.

Pelvic floor muscle performance

The effect of the interventions on pelvic floor muscle performance was determined by PROs (in pelvic floor questionnaire/interview) and by perineometry as a technical measurement. As methodological triangulation (Denzin, 2009), the use of two kinds of measurements with two methods of data collection in the same study and the same research approach should enhance outcome information on the variable pelvic floor muscle performance. Both methods of measurement are considered now.

a) Participant reported pelvic floor muscle outcomes

A patient reported outcome measure is "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else" (U.S. Department of Health and Human Services (Food and Drug Administration), 2009, p. 2). Patient reported outcomes serve to improve the value of research for patients by considering results that matter most to them (International Consortium for Health Outcomes Measurement, 2018b) and have been created in collaboration with patient representatives (International Consortium for Health Outcomes Measurement, 2018a). Pelvic dysfunction is one of the patient reported health and wellbeing [sic] outcomes related to pregnancy and childbirth that the International Consortium for Health Outcomes Measurement (2018c) holds in its compilation. Patient reported outcomes therefore were incorporated into this feasibility trial. For the midwifery focus in this study on a healthy population, the term patient was changed to participant, thus leading to the term participant reported outcomes.

Participant reported pelvic floor muscle performance[67] was measured via a pelvic floor questionnaire (Appendix 20a/b of research protocol [Appendix N]). At the trial planning stage, existing pelvic floor questionnaires[68] and related articles were screened to identify an appropriate questionnaire for the purpose of this trial. Out of the available options, none was considered an optimal choice for this trial because they were either too pathologically oriented, too extensive, did not focus specifically on the muscular aspects of the pelvic floor, or/and asked irrelevant questions; neither could there be found any other appropriate PRO measurement method in the studies screened. Therefore, a questionnaire to determine subjective perception of pelvic floor muscle performance after childbirth was designed specifically for this trial by extracting appropriate questions from the screened literature. It should contain questions for the domains pelvic floor muscle strength, vaginal and anal symptoms, and urinary incontinence.

For the domains pelvic floor muscle strength and vaginal/anal symptoms, the sources for the questions used were an article by Dietz et al. (2012), the pelvic floor and birth questionnaire by Thibault-Gagnon et al. (2014), and Baessler and Kempkensteffen's (2009) German pelvic floor questionnaire for practice and research. Permission was sought from the authors to use questions from their questionnaires and publications. Most of the original questions were (slightly) adapted to suit the purpose of this trial, and response scales were aligned to create a consistent questionnaire design; necessary translations were done by CO. Appendix Z gives details on the origin and adaptations of the questions and answering options used. To ease analysis, the ordinal responses (with four choices each) of the participant reported pelvic floor measurement variables were dichotomised, accepting the loss of information coming with dichotomisation of outcomes (Sankey and Weissfeld, 1998).

Although pelvic floor dysfunction was not the focus of this research, urinary incontinence symptoms were assessed. This was done at trial entry to determine which women would be interested in such a trial and to be able to consider urinary incontinence as a confounding factor. Therefore, the pelvic floor questionnaire also comprised the ICIQ-UI SF (German) (Bristol Urological Institute, 2014b). This is a standardised urinary incontinence questionnaire validated by Avery et al. (2004) in a validated translation by the International Consultation on Incontinence (ICI). There is permission from the ICI to use the questionnaire in its published

[67] Apart from pelvic floor muscle performance, there were other PROs in this trial: adherence, adverse events, qualitative effect answers, and participants' experiences with the interventions and the trial.
[68] These comprised: the Pelvic Floor and Birth Questionnaire by Thibault-Gagnon et al. (2014); the Pelvic Floor Distress Inventory-20 (PFDI-20) short form and Pelvic floor Impact Questionnaire-7 (PFIQ-7) short form from Barber et al. (2005); the German pelvic floor questionnaire by Baessler and Kempkensteffen (2009)/Baessler and Junginger (2011); the International Consultation on Incontinence Modular Questionnaire (ICIQ) (Abrams et al., 2006, Bristol Urological Institute, 2014a); the Kings Health Questionnaire (KHQ) (Cardozo and Kelleher, 1997); Vaizey et al. (1999) for different faecal incontinence grading systems, the Female Sexual Function Index (FSFI) (Bayer AG et al., 2000).

form (Bristol Urological Institute, 2014c). The research plan had been to collect the ICIQ-UI SF sum score only before the intervention for the above reasons. However, it was by mistake also included in the final pelvic floor questionnaire which was realised only at the data analysis stage.

As the pelvic floor questionnaire (except the ICIQ-UI SF part) was self-designed specifically for this trial, no data on its validity and reliability can be provided. However, it was aimed to enhance validity and reliability by using validated items and items from validated questionnaires, by keeping to instructions for good questionnaire design[69] (Burns, 2000, Hulley et al., 2013), and by considering PPI process input by two women (section 5.7). The questionnaire fulfils almost all of the ideal properties of a PRO instrument as summarised by Deshpande et al. (2011), e.g. specificity to the concept being measured, an optimum number of items, or being easy to understand for the intended population. The original validity and reliability characteristics of the different participant reported outcome measure (PROM) items used in the questionnaire are summarised in Table 8.

The questionnaire (in paper form) was completed by each participant during the initial and final study visits. As CO was present, required explanations could be given. After initial experiences with questionnaires being handed back incomplete (e.g. no number noted for pelvic floor muscle strength rating in Question 1 or erroneous boxes ticked), each returned questionnaire was immediately checked for appropriate completion of all items and participants were asked for corrections if applicable.

In addition to these quantitative PROMs in the pelvic floor questionnaire, participants provided qualitative outcome data by answering two open-ended questions in the final interview. These asked about pelvic floor changes since using the intervention and about an ascribed influence on sexual sensations (a more detailed look into sexuality would have made the present trial too voluminous; also, the interest is specifically on aspects of sexuality in connection with pelvic floor muscle performance). The respective methodology is described in the section on participants' experiences and opinion on page 108.

[69] E.g. considering open-ended and closed-ended questions, optimal formatting, or simple concise wording.

Table 8 PROMs' characteristics

Measure	Validity	Reliability
Question 1 Participant reported pelvic floor muscle strength	The percentage scale might be more responsive (and thus valid) than asking via the Likert scale in postintervention Question 2.	Dietz et al. (2012) report a test-retest repeatability with an *ICC*[70] of .75 [.47–.90][a].
Questions 2–9 (preintervention, 2–10 postintervention) Symptom and bothersomeness questions	Questions were taken from existing questionnaires: • Baessler and Kempkensteffen's (2009) validated German pelvic floor questionnaire for practice and research • Pelvic floor and birth questionnaire by Thibault-Gagnon et al. (2014) with limited validity.	Reliability information for source questionnaires: • Self-administered version of Baessler and Kempkensteffen (2009): Cronbach *α*[71] values for internal consistency between .76 and .86 for the different domains (bladder, bowel, prolapse); *κ*[72] values for test-retest reliability between .7 und 1.0 • Thibault-Gagnon et al. (2014): Cronbach *α* values for internal consistency of items within each domain were ranging from .57 to .84 across domains, weighted *κ*[73] values for repeatability between .06 and .89 (median .63)[a].
ICIQ-UI SF (Avery et al., 2004, Abrams et al., 2006)	• Good construct validity • Acceptable convergent validity[74] • Validated German translation	Good reliability

Note: *ICC* = intraclass correlation coefficient.
[a]Values referring to English language version.

[70] The *ICC* is a measure for the reliability of measurements (MedCalc Software bvba, 2018).
[71] Cronbach's *α* as measure of reliability or internal consistency can lie between 0 (minus infinite) and 1 with a value of more than .7 considered acceptable (how2stats, 2015, Statistics How To, 2018, Stats Make Me Cry Consulting, no date).
[72] *κ* as an agreement measure for nominal data can lie between 0 and 1, with .6 or more representing good agreement (Petrie and Sabin, 2009, Bowers et al., 2013).
[73] Weighted *κ* as an agreement measure for ordinal data is interpreted the same way as *κ* (Petrie and Sabin, 2009, Bowers et al., 2013).
[74] Convergent validity is a subtype of construct validity, meaning that "measures that should be related are in reality related" (Trochim, 2006a, www.socialresearchmethods.net/kb/convdisk.htm).

b) Vaginal manometry (perineometry)

To have an intersubjective[75] technical measure to complement the PROMs, an external measurement by assessors was desired. There is no single gold standard measurement tool that tests all aspects of pelvic floor muscle function (Frawley, 2006, Frawley et al., 2006). Of the nine techniques available to measure pelvic floor muscle performance ('Chapter 5: Measurement of pelvic floor muscle function and strength, and pelvic organ prolapse', 2015, Deegan et al., 2018), it was decided to use perineometry as a manometric measurement of intravaginal pressure (Bø, 2015c). It was chosen because it is minimally invasive, the measurement device was thought to be accessible and affordable, assessors were likely to be found, and it was deemed acceptable during discussions with PPI participants. To measure voluntary contraction strength, it is preferable to the alternative digital measurement for its higher reliability (Frawley et al., 2006).

The necessary technical measuring device is called perineometer and was invented and named by Arnold Kegel in 1948 (Kegel, 1948). Today, various perineometer models are in use (Bø, 2015c), of which the Peritron™ PRN09301 (or cat 9300V) with its vaginal sensor by LABORIE (2018, Figure 15) was used in this trial[76]. It has formerly been used as research instrument (e.g. by Reilly et al., 2002, Frawley et al., 2006, Gameiro et al., 2010, Baracho et al., 2012), and among the perineometer models was affordable and could be purchased. Its vaginal probe is 108 mm long with a 55 mm long pressure sensitive zone and of medical grade silicone rubber sheath which was covered with a latex sleeve for each participant. It has to be inserted until 1 cm remains outside of the body so that the pressure sensitive part is at the level of the levator ani approximately 3.5 cm inside the vagina (Bø et al., 1990b, Bø et al., 1990c) as this is the vaginal pressure sensitive zone (Jung et al., 2007). Its diameter[77] is 26-28 mm at insertion and 33 mm when, as recommended by the producer for squeeze pressure measurement (LABORIE, 2012), inflated. With a tube, the vaginal insert is connected to a handheld microprocessor which measures the conveyed pressure.

[75] For a discussion on intersubjectivity as consensual objectivity versus objectivity see Ziman (2000).
[76] Amendment to safety data provided in the research protocol: The Peritron corresponds to Federal Communications Commission (FCC) rules (LABORIE, 2012). The FCC is an “independent U.S. government agency overseen by Congress” and “the federal agency responsible for implementing and enforcing America’s communications law and regulations” (Federal Communications Commission, no date, https://www.fcc.gov/about/overview).
[77] The optimal probe size for perineometry is not known ('Chapter 5: Measurement of pelvic floor muscle function and strength, and pelvic organ prolapse', 2015, Bø, 2015c, Deegan et al., 2018).

Figure 15 The Peritron perineometer

Picture provided by and reprinted with permission of LABORIE.

The different perineometry measures in this trial are explained and justified in Box 4. They were taken with participants in a bent-knee supine position as position with the highest participant acceptance and assessor preference (as most convenient, easiest to standardise and least time-consuming) (Bø and Finckenhagen, 2003, Frawley et al., 2004). In contrast to involuntary reflectory pelvic floor action in real life, perineometry measures a voluntary contraction. This therefore represents an indirect measure of real-life pelvic floor muscle performance (Peschers et al., 2001, Dumoulin, 2004, Bø and Sherburn, 2005, Bø, 2015a, Deegan et al., 2018), which is thought to specifically measure the puborectalis muscle (Jung et al., 2007). Although face and content validiy of pelvic floor muscle pressure measurement is highest in the urethra where the contraction shall exert its required effect, vaginal measurement is preferred as women have a better feeling in the vagina and it is less invasive with a minimal risk for infection (Bø, 2015c).

Vaginal squeeze pressure measurements generally have shown satisfactory reliability (Bø, 2015c). Reliability results of different perineometry measurements with the Peritron device are provided in Appendix AA, together with information on sensitivity. As the measurements can only be valid and reliable when technical details are considered (Bø, 2015c), and to assure these requirements in the present trial, a measurement standard was created and applied (including the verbal instruction, as suggested by Messelink et al. (2005)). Information on the necessary details and the standard are also given in Appendix AA.

Box 4 Measurements taken with vaginal manometry

Vaginal resting pressure (in cm H_2O)
This is vaginal pressure without the participant performing any action (Bø et al., 2013). Its value represents "the passive closure forces [for the urogenital hiatus] from connective tissue and resting muscle tone" (DeLancey et al., 2007, p. 296). DeLancey et al. (2007) and Brækken et al. (2014, p. 118) demonstrate that vaginal resting pressure can be an important marker of ''muscular closing'' of the levator hiatus, and according to Dietz and Shek (2008a), resting pressure might be more important than muscle strength in the etiology of pelvic organ prolapse.
Vaginal squeeze pressure strength (in cm H_2O)
As a maximum force muscle contraction, this represents muscular strength, a component of muscular performance (DiNubile, 1991, Ratamess, 2012). A maximum strength muscle contraction is defined as the maximal force a muscle "can generate during a specific movement pattern at a specified velocity of contraction" (Knuttgen and Kraemer, 1987, p. 7). The participant was asked to squeeze her pelvic floor muscles three times as strong as possible around the vaginal probe. Three repeat contraction measurements should enable the participant to reach her best muscle performance (Bø, 2015c).
Vaginal squeeze pressure endurance (in seconds above 5 cm H_2O)
Muscular endurance as the "ability to sustain performance and resist fatigue" (Ratamess, 2012, p. 11) is another component of muscular performance (DiNubile, 1991). The participant was asked to once squeeze her pelvic floor muscles strongly and as long as possible (Bø, 2015a), up to 10 seconds as this duration is used in clinical judgement (Laycock and Jerwood, 2001). This measurement was operationalised as duration (in seconds) of a contraction sustained above 5 cm H_2O. Of the different methods of endurance operationalisation (Cosner et al., 1991, Wilson and Herbison, 1998, Kerschan-Schindl et al., 2002, Marshall et al., 2002, Hundley et al., 2005, Rahmani and Mohseni-Bandpei, 2011, Sigurdardottir et al., 2011, Friedman et al., 2012, Hilde et al., 2013b, Bø, 2015a), this was chosen because it was the choice available via the Peritron.

At the initial study meeting, participants were informed by CO about the measurement appointment and procedure (information checklist in Appendix AA). They were again explained the procedure just before the measurement by the assessor who also explicitly asked for consent. Precautions were taken to protect the participants' privacy, and medical hygiene guidelines were followed.

Manometric measurements were taken before and after the intervention to collect a maximum of information for the statistical calculations (Hulley et al., 2013). After study entry and the initial study meeting, pelvic floor muscle performance by perineometry was aimed to be measured within 3 weeks to have the measurements as close as possible to the initial PROs data collection (Feasibility criterion 2). At the end of the intervention, it was aimed to

have both measurements within 2 weeks (Feasibility criterion 5) as reversal of muscle training effect (detraining) happens within weeks (DiNubile, 1991, Bø and Aschehoug, 2015).

Two variable operationalisations were done in preparation for the analysis. Vaginal squeeze pressure strength during perineometry was operationalised twofold as 1) the highest of the three contraction measurements (as described e.g. in Frawley et al. (2006), Sigurdardottir et al. (2011), Baracho et al. (2012), Zizzi et al. (2017)) and 2) the mean of the three contraction measurements (as described e.g. in Cosner et al. (1991), Wilson and Herbison (1998), Dias et al. (2011), Friedman et al. (2012), Hilde et al. (2013b)). This operationalisation reflects the fact that both these values can be used to assess muscle performance (Bø, 2015c). Vaginal squeeze pressure endurance was used 1) as measured originally and 2) as endurance (10) trimmed at 10 seconds which reflects the maximum of 10 seconds used to measure muscle performance in pelvic floor assessments (Laycock and Jerwood, 2001, Devreese et al., 2004, Slieker-ten Hove et al., 2009).

Three people in the research team (see Appendix B) helped with outcome assessment. The first nine measurements were done by EH; after this, Sabine Clauss (SC) and Edona Berisha (EB) started contributing, and organising a rota for the measurements was needed. All assessors were given an introduction to the measurement standard and device. They were blinded to intervention allocation and participants were instructed not to disclose their assigned group to the assessors. To communicate the measurement results, an online form was created via formAssembly[78] at City University London into which assessors entered the data.

Harms

A policy of active (participants are asked) and passive (participants report spontaneously) surveillance (Ioannidis et al., 2004) was applied during this feasibility trial to monitor adverse events after uptake of the assigned intervention. An adverse event thereby was defined as "any unfavourable and unintended sign […], symptom, or disease temporarily associated with the use of a medicinal product, whether or not considered related to the medicinal product" (European Medicines Agency, 1995, p. 3). In spite of this definition, only adverse events which could be assumed to potentially be adverse reactions were screened for.

Participants were informed in written (participant information sheet) and oral (at information/consent/initial study meeting) form about potential adverse intervention effects and their warning signs. They were encouraged to contact CO at any time if needed. To check for adverse events, participants were phoned by CO according to the following approximate scheme: four days after the initial perineometry measurement (i.e. after

[78] formAssembly is a web application to design and execute web surveys (Veer West LLC, 2018).

assumedly having started the intervention), one week later, another two weeks later, another three weeks later, and again three weeks later. The content of the calls was standardised (see Appendix T) and the contacts were documented. Adverse events assessment was also part of the final interview.

Adverse events recording comprised type, severity and number of all expected and unexpected adverse events for all participants (Ioannidis et al., 2004). It relied on self-reporting from participants with no medical ascertainment being required. Details considered on potential adverse effects of the interventions can be found in Appendix 8 of the research protocol (Appendix N), and adverse events per group expected at trial start are listed in Appendix 9 of the research protocol.

The described harms assessment in this trial was self-designed, keeping to the mentioned trial guidelines. The assessment seemed appropriate, responsive, sensitive, and specific; construct, content and face validity seemed to be given.

Participant experiences and opinion

The value was emphasised of qualitative data to accompany quantitative research (Craig et al., 2006) and of including qualitative research in trials (O'Cathain et al., 2014, Boeije et al., 2015). Likewise, Kvale (2013) designated interviews as an auxiliary method in conjunction with other methods to (e.g.) ask participants in postexperimental interviews on how they understood the design, and "debriefing" participants and staff on how a trial could be improved was also recommended by Hulley et al. (2013, p. 159).

Therefore, the present feasibility trial also set out to survey participants on their opinion and experiences. At study entry, participants' former experiences with and views on the interventions and the trial were sought; after the intervention period, their perspectives on and experiences with the trial interventions and their trial participation were gathered. Participants' spontaneous remarks were also collected throughout. All this should serve to inform a future trial's design and conduct.

Table 6 and Figure 14 (section 5.4.1) show that the first 18 participants of the experimental and the first 10 participants of the comparison group (together half of the participants) were interviewed by "long" structured interviews before and after the intervention. It originally had been planned to only interview this first half of the participants. However, upon arrival at the last planned interview it became clear that some issues still needed to be clarified in person (and not only online, see below) with each participant to get cardinal information. Therefore, "short" structured interviews before and after the intervention were performed with the remaining 26 (as two withdrawn/excluded) participants, and the schedules for these interviews were created by shortening the long interview schedules.

In particular during the final interviews, a balance had to be found between letting participants express themselves freely in conversational form and completing all items on the interview schedule. Interview recording for documentation and analysis was done by note-taking during the interview and assisted by remembering the conversation at the time of entering data (Kvale, 2013). If items were overlooked during the interview or if uncertainties about the answers arose during transcription, these were clarified for the initial data collection tools at the final interview and for the final data collection tools at the final phone call. Although this could have introduced bias by participants later not answering facts the same way than they would have done at the planned data collection timepoint, it led to a more complete dataset.

From the 28 long final interviews, an online questionnaire was developed to anonymously survey the remaining participants after the intervention and to pilot this anonymous survey questionnaire for the main trial. The anonymous feedback opportunity should prevent social desirability[79] bias. The remaining 26 participants were thus administered this anonymous online survey questionnaire. They were sent its link via a personalised e-mail after the final study meeting or measurement, whichever came later (only the first participant to have the measurement after the visit was sent the link after the visit but before the measurement because the process had not been fully thought out yet); a paper version of the online survey questionnaire, as suggested by PPI participants for participants who might not have Internet access, was not needed. This was done on the same day (in 19 cases) or 1-3 days later (in six cases), only the first questionnaire was sent 6 days later as it had not been finalised before. Participants were reminded of the online survey at the final phone call approximately two weeks later, and by a personalised mail/text message approximately another two weeks later if they had not declared in some way that they already had completed the questionnaire. The online survey was closed on 6 September 2016.

The survey data collection tools were self-designed. As this trial was a preliminary study in preparation for a full trial, the forms were not piloted before the feasibility trial but the feasibility trial should contribute to the evaluation and development of the final interview and online survey schedules. The tools were enhanced by keeping to established guidelines for optimal interview schedule and questionnaire design (Burns, 2000) (see Footnote 69). Interview schedules and the online survey questionnaire were reviewed by two women as part of the PPI process (described in section 5.7).

[79] Tendency of research participants to give responses they believe are socially desirable instead of responses that reflect the truth to provide a better impression of themselves (Oppenheim, 2000).

5.5 Data processing and analysis

The next two sections cover data processing and analysis.

5.5.1 Data processing

Data were collected on paper forms or electronically via online forms created with formAssembly at City University London. The demographic and clinical baseline data collection form, the ICIQ-UI SF and the interview schedules were precoded (see appended data collection forms). The self-designed part of the pelvic floor questionnaire was not precoded to not disturb its appearance, and the electronic perineometry and online survey forms were not precoded as this was technically not possible.

All data collected were entered manually by CO into the computer database with the computer programmes SPSS and Word. Data transfer from paper collection forms into the electronic trial database as soon as possible should enable better remembrance of the meeting content and understanding of the written notes (Kvale, 2013). Correct data entry for quantitative data was assured by four recommended validation rules (Petrie and Sabin, 2009, Stengel, 2010): double checking entered data against the original forms at a later date (all data), date checking (dates), and range checking (numerical data) or scanning by eye (if range checking was not applicable). To ensure accuracy of qualitative data transcription, entered data were checked against the original forms at a later date, as recommended by Oppenheim (2000) and Gibbs (2012).

Data were processed and analysed using the computer programmes Office version 365 (Microsoft, 2018), IBM SPSS Statistics versions 22.0-24.0 (IBM Corporation, 2012/13), R version 3.5.0 (The R Foundation, 2018), and VeraCrypt (instead of planned truecrypt as this programme was not available any more). There was one word document for recruitment analysis and a general SPSS file for all other data. Data cleaning and preparation was done in this SPSS file. Thereafter, descriptive statistics and RR were analysed from this file, whereas all other effect size calculations and some graphs were performed with R. The R analyses were peformed in collaboration with statistician Wolfgang Peter, who also produced the five effect size figures in section 8.2 and the figure in Appendix BB. Qualitative interview data were transferred for analysis from the SPSS file to Excel files.

Statistical online calculators used were the percentage calculator by Lissworx (2016), the CI calculators for proportions by Allto Consulting (2018), StatPages.net (2009) and AusVet (2018), the correlation CI calculator by how2stats (no date), and the F distribution calculator by StatTrek.com (2018).

5.5.2 Data analysis

The methods of analysis for the different data sources and types are described and justified in the next five sections. They comprise the process, management and resources analysis, the clinical data analysis including participants' demographic and clinical baseline data and the preliminary analysis of effect and harms, the analysis of the interview and online survey, with the analysis methods used to assess overall trial feasibility closing the section.

Trial process data

Strengths, weaknesses and peculiarities of each trial process are described, considering input questions on extracting relevant information by Thabane et al. (2010), Tickle-Degnen (2013), and Orsmond and Cohn (2015), partly reproduced in Appendix CC. Adherence is calculated in descriptive form with both the original and the newly developed criteria. Where applicable, the feasibility rate (Box 3) with its 95% CI was calculated and the performance of each feasibility criterion determined after the completed respective data collection.

Trial management and resources data

Human, organisational, and data management issues are described. Applicable questions in relation to study management analysis in a feasibility study from those provided by Thabane et al. (2010), Tickle-Degnen (2013), and Orsmond and Cohn (2015), partly reproduced in Appendix CC, were used to support reflection on and identification of management issues. This process also had an informal component in that aspects of management assessment were completed on the basis of the observed research experience without a specific guiding methodology.

Staff tasks and time, facilities and material needed for all trial processes was calculated from documentation during the trial. Applicable questions in relation to the resources analysis in a feasibility study from those suggested by Thabane et al. (2010), Tickle-Degnen (2013), and Orsmond and Cohn (2015), partly reproduced in Appendix CC, were used to support the respective reflection.

Demographic and clinical data

The clinical analyses concerned participant characteristics per group and the future trial's primary outcome measures effect and harms. As for the above analyses, the questions from Thabane et al. (2010), Tickle-Degnen (2013), and Orsmond and Cohn (2015), partly reproduced in Appendix CC, were used to support and complement feasibility reflection and analysis in addition to the now following statistical analyses.

Participant characteristics

Demographic and clinical baseline data, and motivation, are provided in standard descriptive form for the total of participants and per group. For continuous and discrete data with symmetric and skewed distributions, means with standard deviation (*SD*) and range, or medians with minimum and maximum values and interquartile range (IQR) were calculated; for categorical data, frequencies and percentages in categories were calculated. Birth weight is not only given in continuous but also transformed into dichotomous form of ≤/> 4500 g to determine if and how many participants ever have given birth vaginally to a heavy newborn[80]. Previous urinary incontinence and all characteristics of past births were coded into dichotomous form to determine if the events of interest ever happened.

Effect

Descriptive exploratory analysis of the outcome results was used to verify the quality of the data (Appendix BB). Due to the feasibility design of this study, the effect analysis is also exploratory and the calculated effect results are preliminary. As not planned as primary outcome of a future full trial, and as their effect size would not contribute to informing a future trial, the dichotomised PRO symptom and bothersomeness question results are compared in descriptive form only. As the postintervention scores of the ICIQ-UI SF were available unexpectedly (see page 102), it was decided to benefit from the situation and also analyse urinary incontinence descriptively. The descriptive results are reported per group and by the same summary measures as the participant characteristics.

The effect analysis deviated from the research protocol due to the PhD learning experience and subsequent statistical consulting advice. This revealed that mixed modeling (instead of the planned repeated measures analysis of variance [ANOVA]) is the most appropriate analysis method for repeated measures of outcome variables, particularly for the three repeats that were taken for contraction strength. However, as a mixed modeling analysis would not have been proportionate to the respective research question of this trial, and as the question about preliminary differences between the groups' change scores could be answered with the simpler technique of the *t*-statistic, this was preferred. Contrary to the research protocol, no sensitivity analysis was performed for potential confounding variables and missing data as this would not have contributed to informing the future definitive trial (Eldridge et al., 2016a).

For the PRO Question 1 and for perineometry, results are reported in descriptive form, and effect sizes in the form of change score differences are calculated for between groups comparisons. The questions guiding these effect size analyses were:

[80] Definition of high birth weight according to Salvatore et al. (2017).

How much does the change in outcome values from pre- to postintervention measurement differ between the intervention groups?

The calculated effect sizes are given in simple (unstandardised) form with their 95% CIs, thus presenting magnitude and direction of point estimates and their variances (Thompson, 2002, Durlak, 2009). The calculation was performed via the *t*-statistic for independent samples. Welch's modification was used to adjust for unequal group size and the small sample (how2stats, 2014). For the PRO postintervention Question 2, the effect size analysis question was:

How much does the outcome value differ between the intervention groups?

For the between-groups comparison of this (dichotomised) variable measured only after the intervention, the risk ratio (RR) with its 95% CI was chosen in accordance with Schmidt and Kohlmann (2008) and Cochrane review practice (Herbison and Dean, 2013).

To maintain the effect of randomisation, an intention to treat (ITT) analysis comparing all randomised participants by assigned treatment is the recommended RCT analysis strategy (Hollis and Campbell, 1999, Higgins et al., 2011b, Meinert, 2012, Joshi et al., 2013). Due to the exclusion of one participant from the trial after randomisation, the analysis is labelled modified ITT (mITT, Gravel et al., 2007, Gupta, 2011). The mITT analysis was performed as available case analysis, including all participants for whom outcome values were obtained (Higgins et al., 2011b).

In an ITT analysis, protocol deviations can dilute the treatment effect (Montori and Guyatt, 2001). In contrast to an ITT analysis, a per protocol (PP) analysis only includes participants who adhered to the research protocol, whereby an ITT analysis is more conservative and less likely to show an effect than a PP analysis (Thabane et al., 2013). Therefore, a complementary PP analysis for the potential future primary clinical outcomes compared the groups on the basis of intervention adherence and trial completion to test the robustness of the mITT analysis.

Harms

Like the effect analyses, the harms analysis is exploratory and the results are preliminary. Leon et al. (2011) stress that a feasibility study primarily develops the adverse event reporting system, and Arnold et al. (2009, p. S73) speak of screening for “potential harm”. Analysis of harms in trials is recommended to usually be descriptive (Ioannidis et al., 2004). As only the participants having used an intervention can experience side effects that could be caused by it (Higgins et al., 2011a, 'Intention to treat analysis and per protocol analysis: complementary information', 2012), a PP instead of an ITT analysis is used in this case.

Expected and unexpected adverse events and treatment emergent adverse events are reported for all participants with type and number according to interventions, and presented in group-specific rates with 95% CIs (Leon et al., 2011). If applicable and available, recurrence is included; however, as adverse effects were collected at different forms and time points, it can from hindsight only be determined whether they occurred or were mentioned more than once with certainty for the outcome vulvovaginal symptoms, but not for muscle soreness, discomfort/pain, bleeding, and other adverse events. Nonphysiologic bleeding was only explored in the experimental group as this is not a known or expected side effect of pelvic floor muscle training.

If adverse events were identified, the criterion for approximate normality (*CAN*, Formula 3 in Appendix G) was calculated to decide on the appropriate method to determine the 95% CI (Jovanovic and Zalenski, 1997). A *CAN* > 5 allowed calculation of approximate CIs, a *CAN* ≤ 5 (few adverse events) suggested the calculation of exact CIs. If no adverse event was detected in a group, the rule of three (Hanley and Lippman-Hand, 1983, Jovanovic and Levy, 1997, Jovanovic and Zalenski, 1997) was applied to estimate the upper bound of the 95% CI. This rule states that 3/(n+1) is a "very good approximation of the exact upper 95% confidence limit" for binomial probability when no events occur in *n* independent trials (Jovanovic and Zalenski, 1997, p. 303). When counting adverse events, only the upper limit as the worst case scenario (highest rate of adverse event) is of interest, while the lower bound of the 95% CI is not of interest and set as 0 (no adverse event). The application of the rule of three had not been planned in the research protocol but was discovered with subsequent reading.

Survey data

The interviews and the online survey produced factual information in the form of quantitative and qualitative (answers to open-ended questions) data which were analysed after termination of data collection. For quantitative data this was done in SPSS or Excel and results are reported by frequencies and percentages or by means. Qualitative data were analysed by content analysis, "a technique for a systematic quantitative description of the manifest content of communication" (Kvale, 2013, p. 105). This followed the principles of processing qualitative survey data given by Oppenheim (2000). In this data driven coding (Gibbs, 2012), coding frames with answering categories were created from sifting through the answers received for each open question. The data thus condensed to a few categories were amenable to statistics in that categorised responses could be counted (Gibbs, 2012). Qualitative data so became quantifiable and then followed the processing, analysis and reporting of quantitative data.

Validity and reliability of the qualitative analysis and thus of the interview knowledge produced was enhanced by validation techniques. Accuracy and consistency of categorising was checked by controlling whether all data had been attributed to categories and by constant comparison (looking for consistency/differences/variations) of coded text within and between categories (Gibbs, 2012). If necessary, coding was revised until all data were coded. If answers could have been attributed to more than one category, the best fitting category was chosen. Answers to a question which in fact answered another question were attributed to the question they answered.

Integration of feasibility results

The feasibility analysis was inspired by various sources. Most important was the calculation of feasibility rates (Thabane et al., 2010, Abbott, 2014, Whitehead et al., 2014, Eldridge et al., 2016a) as planned in the research protocol. During the analysis stage, further methods were added. One was the consideration of questions to be asked and answered in feasibility studies as suggested by Thabane et al. (2010), Tickle-Degnen (2013) and Orsmond and Cohn (2015). Shanyinde et al.'s (2011) list of 14 methodological issues to be evaluated in feasibility/pilot studies was used to support reflection about the issues addressed in this feasibility trial, together with the questions and answering scheme derived from this list by Bugge et al. (2013). Bugge et al.'s (2013) own framework to analyse feasibility studies, a process to support robust and systematic decision-making in moving from a feasibility trial to a full RCT, served to identify and appraise feasibility problems and their potential solutions. Finally, the CONSORT 2010 guideline extension for reporting of randomised pilot and feasibility trials (Eldridge et al., 2016a) was consulted. Although there is room for subjective interpretation, the methods used seemed appropriate, sensitive, and specific and to have content, face and predictive validity[81].

Integration of quantitative and qualitative results

There is no explicit quantitative or qualitative research question or data collection area in this feasibility trial (the survey e.g. is conducted with interview schedules and online questionnaires which all contain questions leading to quantitative and qualitative results). Instead, within data collection for the areas trial processes, trial management, clinical outcomes and participants' voices on the trial, both quantitative and qualitative data were collected alongside each other.

The guiding framework used to organise this report are the feasibility areas according to Thabane et al. (2010). Therefore, to inform the findings of this feasibility trial, the combination of the collected and analysed quantitative and qualitative data for presentation of results was

[81] Predictive validity is the "ability of the measurement to predict an outcome" (Hulley et al., 2013, p. 39).

organised according to their contribution to each aspect of a feasibility analysis following Thabane et al.'s (2010) framework; e.g. both qualitative and quantitative clinical results are synthesised in the results chapter on clinical results.

5.6 Ethical issues

Ethical issues in this feasibility trial concerned and concern adherence to relevant ethical principles in study conduct, ethical approval, trial registration and result dissemination.

5.6.1 Relevant ethical principles and study approval

According to the World Medical Association (2018) and the Good Clinical Practice (GCP) guidelines (National Institute for Health Research Clinical Research Network (NIHR CRN), 2016), relevant ethical principles in studies with humans are: doing no harm, participants' right to autonomy, and information and data protection. Health and safety concerns also need to be considered for the researcher (City, University of London, 2018). In relation to this feasibility trial, these principles were specified in the research protocol (Appendix 8 of Appendix N), together with information on how it was planned to address them. Safety/harm of the interventions was considered in particular in Appendix 9 of the research protocol (Appendix N) on adverse events monitoring, recording and reporting. As potential adverse effects were expected to be minor, there were no harms-related trial stopping rules. Once approved, the trial was conducted in accordance with the research protocol and the applicable ethical guidelines. Ethical challenges coming up and dealt with during the trial are reported in the process results section.

Ethical approval was sought and obtained from the Research Ethics Committee of the Medical University of Vienna (lead committee) and City University London Senate Research Ethics Committee (approval documents are provided in Appendix DD). An annual progress report and request for extension was submitted in Vienna in September 2015 and October 2016, and a summary final report will be sent to both Research Ethics Committees.

5.6.2 Registration, reporting and dissemination

After ethics approval, the trial was registered at clinicaltrials.gov (Oblasser, 2015) and the research protocol published in the *Journal of Advanced Nursing* (Oblasser et al., 2016). Until the time of submission of this thesis, this publication and 11 presentations resulted out of the empirical part of this PhD and are listed in Appendix C (corresponding information for the systematic review see section 3.4 and Appendix C). Continuing reporting and dissemination of the knowledge gained is planned via submitting the feasibility trial results to clinicaltrials.gov, public availability of this thesis, further presentations at professional

conferences and publications in appropriate journals. (PPI) Participants[82], recruitment supporters and other study helpers will be sent a (lay) summary of trial results as is recommended (Chalmers, 1995, Partridge and Winer, 2002), and oral presentations will be offered to recruitment supporting sites.

5.7 Public and patient involvement

Public and patient involvement in clinical research is recommended by the National Institute for Health Research (no date-b). It "refers to an active partnership between patients and/or members of the public and researchers" and "can help to make health research more relevant to the needs of patients, carers and service users" and provide "alternative views from those of the research team" (National Institute for Health Research, no date-b, p. 5 and 6).

As the aim of PPI in this study was to make the research more useful to women and enhance the success of the feasibility trial, PPI took place at the level of *consultation*[83] (Entwistle et al., 1998, Boote et al., 2006). Six women having given birth within the last few years were consulted in the development phase of the research protocol for this feasibility trial. The ethnic background of these women was Somali, British White, and Austrian White. The women in the UK were recruited at a Maternity Service Liaison Committee (MSLC) and a Social Action for Health parents' meeting in the London Borough of Tower Hamlets in February and May 2014. At the MSLC meeting (Figure 16), the audience was introduced by CO to the purpose of the PPI work and interested women could leave their contact details; the networking in the Social Action for Health meeting was more informal (the information sheet used in both meetings is enclosed as Appendix EE). In Austria, two clients from CO's past midwifery practice were contacted and agreed to contribute, one of them introduced a further PPI participant.

As it was difficult to find dates for joint meetings, individual meetings with the contributing women were held between March and July 2014, at public places or at women's homes, and lasted one to two hours. Research protocol issues in question at the time of each meeting were discussed with the women, such as where best to recruit participants, their potential questions and worries, number and kind/place of study contacts, or acceptability of the experimental intervention and outcome measurements. Apart from many other contributions to the just named areas, the PPI participants' suggestions included: to articulate the relevance of the research to potential participants, that religious women might not want to use a vaginal ball, to ask a female person for pelvic floor muscle measurement, or to also

[82] Participants could tick in the consent form whether they would like to be sent the trial results–which all did.
[83] The other levels being collaboration and user-control (Boote et al., 2010).

prepare a paper version of the online survey questionnaire as some participants might not have Internet access. Two women checked the appropriateness and comprehensibility of the trial documents for participants.

Figure 16 Patient and public involvement in practice

Picture printed with permission of Hana Xassan.

Summary

The feasibility of the future full RCT is determined for trial processes, management, resources, and by the preliminary clinical trial results, complemented by a survey on participants' experiences with and opinion on the interventions and the trial. Five feasibility criteria were set.

A nonprobability sampling method with different recruitment forms and paths was used to recruit the calculated sample size of 56 participants. The experimental intervention was hold and vibration use of vibrating vaginal balls, the comparison intervention pelvic floor muscle training "at home", and treatment was assigned by blocked randomisation with uneven blocks in a 2:1 ratio. Variables were collected for the feasibility outcomes (processes, management, resources), pelvic floor muscle performance, and participants' experiences with and opinion on the interventions and trial with various data collection methods: trial process documentation and informal observation, questionnaires, interviews, adherence

charts, phone calls, participants' self-reports, and perineometry. Analyses corresponded to the outcomes and comprised statistics, resources calculation, descriptive trial process and management analysis, qualitative content analysis, and overarching feasibility analysis.

Patient and public involvement supported the development of the feasibility trial protocol. Ethical approval was gained by the ethics committees of the Medical University of Vienna and City University London, after which the trial was registered at clinicaltrials.gov. The following four chapters present the results obtained from the described empirical research.

6 RESULTS I – TRIAL PROCESSES

This and the following two results chapters are, like the data analysis section of the methods chapter, organised according to Thabane et al.'s (2010) areas of reasons for conducting pilot/feasibility studies: trial processes, management and resources, and clinical outcomes. They are complemented by a fourth results chapter on participants' experiences with and opinion on trial participation.

This first of the four results chapters looks at the trial processes' aspects and by this contributes to the overall research question on how best to prepare and perform a full RCT on the effectiveness of vibrating vaginal balls to improve pelvic floor muscle performance after childbirth. It further contributes to the research questions on harms of the experimental intervention and on participants' perspectives on and experiences with the interventions and the trial.

It presents the findings of the trial processes according to the logical order of the research process. The process of recruitment is considered first, followed by the sample description and participant flow. Further process results cover the trial forms, the data collection, and the trial interventions.

6.1 Recruitment

The following sections describe the recruitment in terms of success and failure and in relation to the selection criteria.

6.1.1 Recruitment success and failure

The recruitment process for each active and passive recruitment path is described in detail in Appendix M, which also summarises efforts and success to recruit participants via the different paths. This shows that 10 of 31 approached obstetricians and 11 of 21 approached midwives supported recruitment, and that 41 women were recruited via health professionals and 15 via recruitment sheet/online text or word-of-mouth. There were large differences in the number of potential participants contributed by each professional.

When calculating recruitment rates, it can be differentiated between form of information about the trial (active or passive as described in methods chapter) and form of contact (active by CO calling potential participants, or passive by waiting for their call). The respective recruitment rates are

- 45.1% (41/91) for women who were actively informed by their midwife or obstetrician,
- 34.9% (15/43) for women who had somewhere found the information about the study or heard of it,
- 45.1% (41/91) for women who were contacted by the researcher, and
- 34.9% (15/43) for interested women contacting the researcher.
- The overall recruitment rate was 41.8% (56/134).

All these recruitment rates must be differentiated from the recruitment rate out of *eligible* women asked for in Feasibility criterion 1. For active recruitment, this rate might be calculated at two levels: that of all women eligible after childbirth at the recruiting professionals' level, or that of all women eligible after information by professionals at the researcher recruitment level. It was originally intended to calculate this rate at the professionals' level but this did not prove feasible as it would have demanded too much cooperation in terms of attention, time, and documentation effort from the professionals. However, the professionals were asked for feedback on their recruitment success and failure at the end of recruitment, and their answers are incorporated in Box 5.

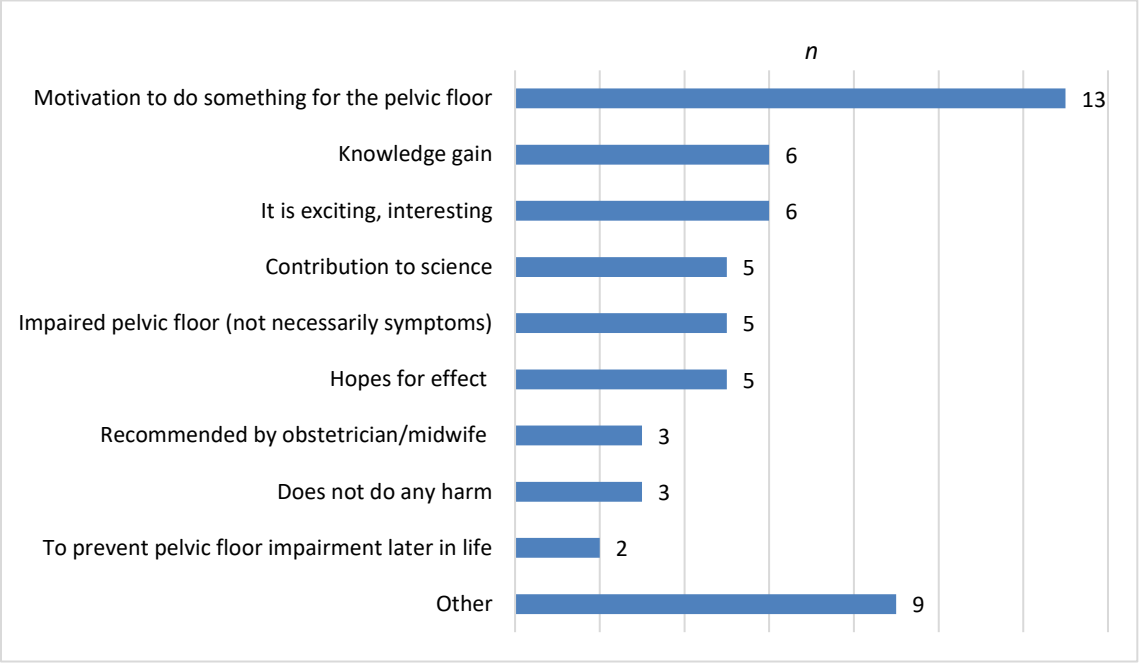
At the researcher level, the recruitment rate out of eligible women was calculated for active, passive and overall recruitment. An overall calculation shows that from 134 women who were named to or contacted CO, 18 women were not eligible. This leaves 116 as possibly eligible as those who were not checked for may have been eligible or not. If the overall recruitment rate is calculated with 116 as the population (assuming all being eligible), it is 48.3% (56/116, 95% CI [39.2, 57.4]); with fewer women assumed eligible, the calculated rate would be higher. In considering active recruitment by form of information, the rate would at least be 49.4% (40/81), the same calculation for passive recruitment gives a recruitment rate of 45.7% (16/35). This result would fulfil Feasibility criterion 1 which asks for a rate of ≥10% out of eligible women; the criterion however considered the basic professional level while the rate is calculated at the researcher level.

Women's answers for their motivation to participate in the trial are categorised in Figure 17, with the most frequently named reason being the "motivation to do something for the pelvic floor". The nine answers that could not be categorised comprise[84]: "because I plan to control the pelvic floor", "I have time", "to be more prepared for a third baby", "I am generally very interested in the body", "I want to find back to my body after birth", "I like exercising, it´s healthy", "because it is a topic which is important for many and is not talked about", "because I have heard that it is important", "widening of consciousness by first birth–a new body part,

[84] German as the author's first language was the language of communication with all except two participants for whom it was English. Even when the participants' anwers were translated into English (all translations by CO), they are put into quotation marks to point out direct quotations.

and one can consciously influence it", "because I want to support you with the research question", and "measurement is interesting". Box 5 on the contrary presents the reasons for recruitment failure, split into the recruitment levels: managerial, individual professional, researcher, and potential participants.

Figure 17 Participants' reasons for participating (*N* = 28, multiple answers possible)



Box 5 Recruitment failure per recruitment level

Managerial level
In London, managerial NHS (National Health Service) contact was unavailable and access to the field thus could not be gained–this was the reason why the trial could not be performed in the UK as originally intended. When planning the trial at the AKH Vienna, it had been considered to include nursing staff on the postnatal ward (no midwives were working there) in recruitment. At a meeting with the responsible nurse manager, she clarified that nurses would not support recruitment due to work constraints (already so many forms to distribute and explain at discharge). After this decline, medical doctors at the postnatal ward of the AKH Vienna were approached. After recruitment via medical doctors at the postnatal ward of the AKH Vienna did not seem feasible (see recruitment level below), parenting centres of the City of Vienna were approached as an alternative. Recruitment sheets were laid out at the parenting centre in Vienna's 21st district. However, during the process of approaching further centres, the author was informed that it was not allowed to private persons to distribute information material in the centres, and therefore further access could not be gained.
Individual professional level
When introducing the study to the medical doctors at the postnatal ward of the AKH Vienna attending a staff morning meeting, the question was raised whether authorship of a publication could be obtained (this issue had already been raised in the initial negotiations about this recruitment possibility). It seemed that one of the more senior doctors might then have taken on responsibility and collaborated. As there was already a research team involved in doing this trial (see Appendix B), CO declined to take aboard other authors who had hitherto not participated in trial development, the more as the other recruitment routes seemed to work. It was nevertheless agreed that junior medical staff would distribute participant information and consent forms to women at discharge and note contact details; this however did not prove feasible either. Seven of the 21 contacted community midwives did not want to support recruitment. Their reasons were: not providing (enough) postpartum care at the moment (*n* = 4), no time (*n* = 1), all clients going to pelvic floor coaching after birth (*n* = 1), and recruitment documentation would ask too much effort (*n* = 1). Fourteen of the 31 contacted obstetricians did not reply to the e-mail asking for recruitment support. The reasons given by four obstetricians for not supporting recruitment were: no time (*n* = 1), too few pregnant clients (*n* = 1), only has private[85] patients who do not like to participate in studies (*n* = 1), and is sending all postpartum women to physiotherapist (*n* = 1).

[85] Private healthcare clients pay the health professional on their own and later are reimbursed (part of) the fees by their health insurance.

Researcher level
Reasons for nonrecruitment (per most appropriate denominator): – Recruitment complete (8/134 contacts [6.0%]) – Could not be reached (at all: 5/91 contacted [5.5%]; for a second call: 2/91 contacted [2.2%]) – Did not get back to researcher although agreed upon (6.2% [8/129 spoken to]) – Declined participation (28.7% [37/129 spoken to]) – Not eligible (24.3% [18/74 checked for]) Reasons for noneligibility (*n* = 18, first reason identified): – Over 6 months post partum (*n* = 4) – Currently enrolled in pelvic floor muscle training (*n* = 4) – Caesarian section (*n* = 3) – CO recommended therapy for urinary incontinence symptoms (*n* = 3) – Preterm birth (*n* = 2) – Ill baby (*n* = 1) – Planning to get pregnant soon (*n* = 1)

Potential participant level

Reasons given to CO by 34 women who declined participation at phone (multiple answers possible, reasons for three further women unknown):

- Does not want to use a vaginal ball or be randomised into experimental group (*n* = 10)
- Time constraints (*n* = 5)
- Too much effort to travel to measurement (*n* = 5)
- Exercising at home (included Yoga or walking; *n* = 5)
- Concerns about adherence (*n* = 4)
- Other (*n* = 10):
  - Will move away, is living too far away (*n* = 2)
  - Is away often or for a longer period (*n* = 2)
  - Single mother (*n* = 1)
  - Does not want vaginal measurement (*n* = 1)
  - Has no pelvic floor problems, participation would be undue stress (*n* = 1)
  - Has no pelvic floor problems, does not want to occupy a study place which could be useful to someone else (*n* = 1)
  - Does not want to go to the hospital (for measurement) with the baby (*n* = 1)
  - Is back to work (*n* = 1)

The recruiting professionals confirmed these categories when reporting the reasons they had been given by women not interested in participation; one private (see Footnote 85) obstetrician communicated that her/his patients did not want to participate in the trial. Only two women did not agree to have their phone number communicated to CO.

In two cases, the information/consent/initial study meeting had been arranged, and when presenting the trial and the ball to be used, the women declined participation. In the first case, the woman suffered from genital scar tissue after suturing at birth and could not imagine using the ball. In the second case, it was the also present husband who first stated this would not be an intervention for his wife; the woman then agreed with him but did not give an explicit reason.

6.1.2 Eligibility criteria

The inclusion criteria in general proved appropriate and feasible. A suggested minor change identified during the trial is to add the exclusion criterion *Status post pelvic floor surgery*. For the criterion *Excluding women currently enrolled in pelvic floor muscle training with physiotherapist, midwife or fitness trainer*, the wording should be adapted to the research reality which also excluded women *planning to enrol in such training during the trial period*. The criterion should also specify to only exclude women enrolled in *individual* training with professional; enrolment in postnatal exercise classes (which have pelvic floor muscle exercises as an integral part) should not be an exclusion criterion any more as going to a class once a week is unlikely to influence muscle performance.

During the trial, a change arose from one potential participant who declined the routine postnatal six weeks check as she considered herself in good health and not in need for such an examination. This led to a review of the respective inclusion criterion and a change (in italics) to the following wording: “Six weeks postpartum check by obstetrician *or other appropriate professional* performed and woman discharged from postpartum care with diagnostic findings appropriate to this period after childbirth; *exceptionally also women who decline this check as they feel in good health*”. This was approved by both ethics committees.

The exclusion criterion *Retention of ball is impossible* proved difficult to determine and apply. The trial protocol had planned that the respective participants use the ball once after inclusion and before the initial perineometry measurement to see whether they can hold it inside the vagina, and to exclude them if they were unable to do so. However, the inegilibility of the one participant concerned who could never hold the ball for more than between three and 15 minutes only became clear after a process of continuously trying, some phone conversations to discuss suggestions to enhance the situation, and after more than two months into the intervention period. According to the criterion, she should have been excluded from the study upon this finding; however, to profit from her experiences and as this became clear so late, it was decided to keep her in the trial.

With the case of candidiasis/cystitis during the trial (see section 8.3), a discussion arose around the exclusion criterion *Termination of participation will be recommended to a participant when there is vaginal infection* which had been planned for the experimental group. As an exclusion on these grounds in the experimental group would introduce

performance bias[86] because participants in the comparison group would not be excluded for this reason, this criterion needs to be deleted in a full trial.

6.2 The sample

Although the sample was planned to be hospital- and community-based, it ended up as community-based sample with only one participant being recruited from the hospital setting. In the following, the participants are described in terms of their demographic and clinical characteristics measured at baseline, and in terms of their experiences with and opinion on the interventions at trial start.

6.2.1 Demographic and baseline clinical characteristics

Table 9 shows the demographic and baseline clinical participant characteristics. Although one participant was excluded after randomisation, the sample description shows all recruited participants as the aim in this process section is to determine feasibility issues with respect to the women getting enrolled in such a trial.

All participants were socially well situated (assessed by being in a partnership and having a professional education). Except for one Chinese woman, the sample is Caucasian only, but it is multicultural with participants coming from the following cultural backgrounds: Austrian (*n* = 40), German (*n* = 5), Hungarian (*n* = 2), and one each from a Czech, Polish, Russian, German/Russian, Slovenian, Serbian, Greek and Australian background. All participants had a singleton birth at term and all except one were within 6 months post partum at intervention start. The one exception surpassed this limit by 5 days as timely measurement could not be organised for her.

In the ICIQ-UI SF, 21 participants (37.5%) reported urinary incontinence symptoms at trial start, with their sum scores ranging from 3-14 (median 6, mode 5.5), and the incontinence forms presented in Figure 25 (page 162).

[86] Systematic differences in exposure to other factors apart from the intervention of interets (The Cochrane Collaboration, 2018).

Table 9 Demographic and baseline clinical participant characteristics (*N* = 56)

Characteristic		Value
Highest completed education *n* (%)	University degree High school Vocational middle school or apprenticeship	42 (75.0) 9 (16.1) 5 (8.9)
Occupational status *n* (%)	Paid work (maternity leave) Housewife Student	48 (85.7) 7 (12.5) 1 (1.8)
Age (years)	Mean (*SD*) Range	33.3 (4.4) 21-41
BMI[a]	Mean (*SD*) Range	22.4 (2.6) 17.9-32.8
Parity *n* (%)	I II III/IV	25 (44.6) 28 (50.0) 3 (5.4)
Completed weeks post partum at inclusion	Mean (*SD*) Range Median (min, max)	12.5 (4.9) 6-24 10.5 (6, 24)
Mode of birth *n* (%)	Spontaneous Ventouse	51 (91.1) 5 (8.9)
Birth injury *n* (%)	Episiotomy 2nd degree perineal tear None	4 (7.1) 5 (8.9) 14 (25.0)
Birth weight newborn (g) *n* (%)	Mean (*SD*) Range > 4500	3399 (398) 2620-4530 1 (1.8)
Past mode of birth *n* (%)	Ever ventouse Ever caesarean section	1 (1.8) 5 (8.9)
Past birth injury *n* (%)	Ever episiotomy Ever 2nd degree perineal tear Ever 3rd degree perineal tear	3 (5.4) 7 (12.5) 0 (0)
Weight newborn ever > 4500g *n* (%)		1 (1.8)
Breastfeeding at study entry *n* (%)		54 (96.4)
Cigarette smoking *n* (%)		2 (3.6)
Any previous urinary incontinence *n* (%)		32 (57.1)
ICIQ-UI SF sum score[b]	0 *n* (%) 3-9 *n* (%) 13-14 *n* (%) Median (min, max) IQR	35 (62.5) 17 (30.4) 4 (7.1) 0 (0, 14) 5

Note. SD = standard deviation; IQR = interquartile range; min = minimum, max = maximum.
[a]Normal weight BMI range 18.5-24.9 (World Health Organization, 2018a).
[b]Maximum = 21.

6.2.2 Participants' experience with and opinion on the interventions at trial start

Of the 56 participants, 50 (89.3%) had ever performed pelvic floor muscle exercises. Of these 50, 21 (42.0%) had attended a postpartum exercise class, seven (14.0%) had performed pelvic floor muscle exercises in a childbirth education class, five (10.0%) had been instructed individually by a midwife or physiotherapist post partum, and 15 (30.0%) had exercised only individually at home (categories mutually exclusive, multiple answers categorised by first to last named here). Accordingly, participants' sources of knowledge about pelvic floor muscle exercises were physiotherapists, midwives, fitness or Yoga classes, the World Wide Web, and books; two of the participants were midwives themselves, one of them was also a Yoga teacher for childbearing women, and three were medical doctors. When asked about pelvic floor muscle exercises, techniques other than repeated contraction exercises were mentioned by 14 participants (of 50, 28.0%) and comprised Pilates (Pilates Foundation, 2018), Yoga, Cantienica® (CANTIENICA AG, 2018), Kanga (*Kanga Training*, no date), and use of a vaginal ball.

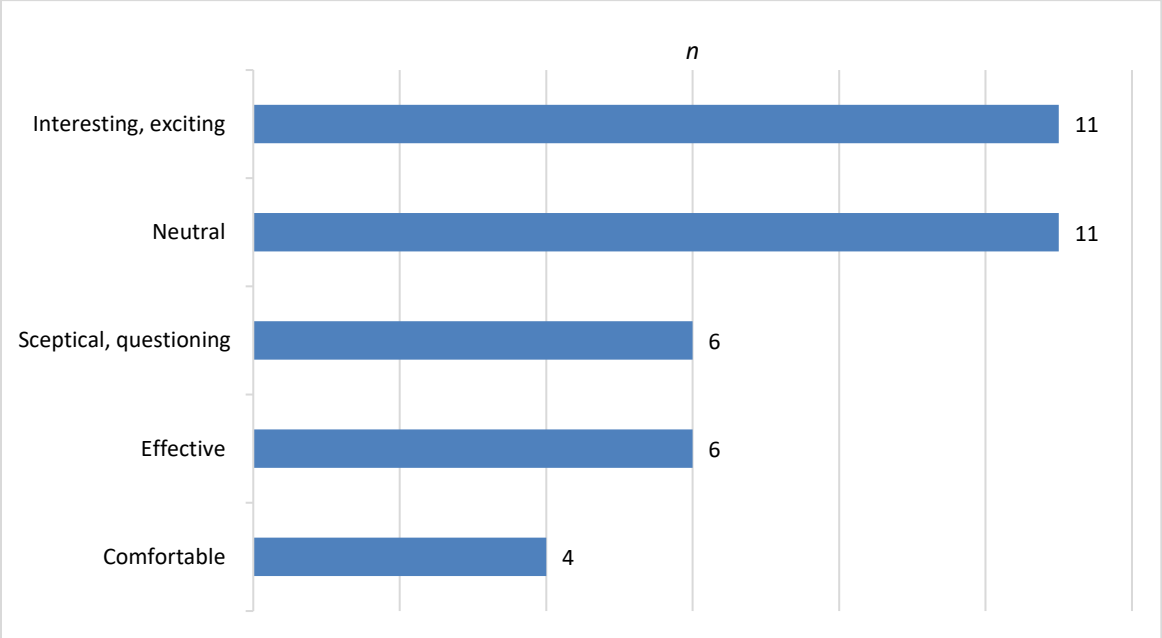
After the most recent birth, 47 participants (of 56, 84.0%) had been given written and/or oral information about pelvic floor muscle exercises, eight (14.3%) said they had not been given any information, and one (1.8%) did not know but remembered exercises from antenatal childbirth education classes. A closer examination of 31 participants' information material revealed that the nature of these exercises always was gentle as aimed at the early postpartum period.

Of all participants, 13 (23.2%) had not exercised at all after the reference birth. Of the 43 who had performed pelvic floor muscle exercises, seven (16.3%) had begun with sessions which decreased over time, and 17 (39.5%) had exercised (very) rarely and on an irregular basis. Nineteen (44.2%) exercised at least two to three times a week at trial start. The nature of the exercises varied and was e.g. 50 contractions throughout the day or exercising twice a week with break-offs.

Asked about their opinion on pelvic floor muscle exercises, 21 participants answered that these are important and useful, five mentioned the preventive aspect, seven commented on motivational issues (time, no problem–no training), three found the pelvic floor muscles difficult to identify or were unsure whether they were exercising correctly, and three found the exercises boring, cumbersome, tingly, or old fashioned. Two participants commented that pelvic floor muscle exercises would be the beginning of re-starting regular exercising after birth and can be practised everywhere, and two expressed hope about their effectiveness.

Two thirds (38/56, 67.9%) of the participants had heard or read of vaginal balls before hearing about this trial, and seven (of 56, 12.5%) had already used such (a) ball(s). One had bought cones before but had never used them. The participants' opinion on the balls at trial start, ranging from positive to sceptical, is summarised in Figure 18.

Figure 18 Participants' opinion on vibrating vaginal balls at trial start (*N* = 28, multiple answers possible)



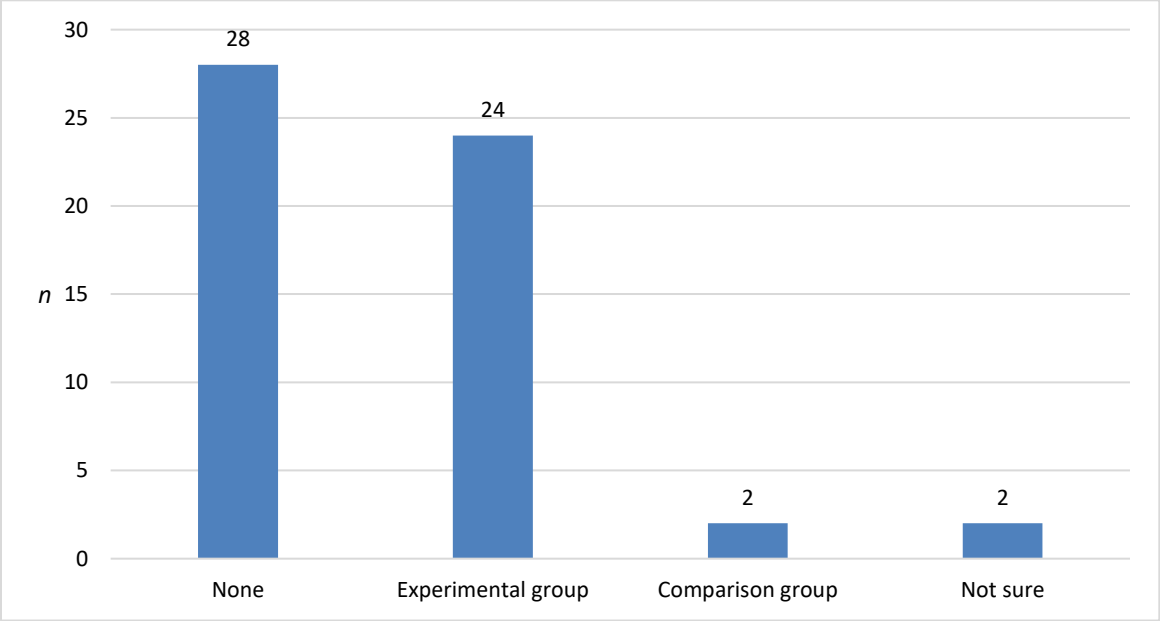
6.3 Randomisation

Randomisation ran smoothly technically and in terms of acceptability. By following the research protocol, randomisation for two participants took place before their ineligibility could be spotted. For one participant, her inability to consciously contract the pelvic floor was identified when checked, as planned, at the perineometry appointment which then led to her exclusion according to the exclusion criterion *Inability to perform the proposed procedures*. For the other participant, this concerned the exclusion criterion *Retention of ball is impossible* and was described on page 126.

Preference for a trial group was stated by 26 (of 56, 46.4%) of the participants, by 24 (42.9%) for the experimental group and two (3.6%) for the comparison group (Figure 19). Of the 26 participants with a group preference, 17 (65.4%) got allocated to their preference group and nine (34.6%) to the nonpreference group.

When asked to comment on their allocation, 31 (83.8%) of the 37 participants allocated to the experimental group gave a positive and six (16.2%) a neutral statement; no negative comments were given. Of the 19 allocated to the comparison group, three (15.8%) gave a positive, 10 (52.6%) a neutral and six (31.6%) a rather negative statement. One participant in the comparison group was happy that in her group "nothing can happen in terms of [intervention] risks" (ID 48), and one was relieved that the training consisted of pelvic floor contractions only without accessory exercises.

Figure 19 Participants' preferences regarding group allocation (*N* = 56)



On a scale from 0 to 10, those allocated to the experimental group rated their motivation to do the intervention for 12 weeks as 9.1 (*SD* 1.2, range 5-10), those allocated to the comparison group as 7.9 (*SD* 2.0, range 2-10). Those allocated to their desired group (*n* = 16) rated their motivation to do the intervention for 12 weeks as 9.4 (*SD* 1.1, range 7-10), those allocated to the nondesired group (*n* = 10) as 7.3 (*SD* 2.3, range 2-10). Table 10 shows the motivation results for desired allocation split by groups.

Table 10 Participants' motivation by intervention group and desired allocation

Experimental group (*n* = 37)			**Comparison group** (*n* = 19)	
n	Mean (*SD*, range)		*n*	Mean (*SD*, range)
21	9.0 (1.3, 5-10)	**No desired group**	9	8.6 (1.2, 7-10)
15	9.3 (1.1, 7-10)	**Allocated to desired group**	1	10
1	8	**Allocated to nondesired group**	9	7.2 (2.4, 2-10)

At the postintervention online survey, all 14 respondents in the experimental group stated to have been happy with their group allocation. When asked for their reasons, seven stated they had wanted to try the ball, four found using the ball more practical and leading to higher adherence, and one each thought pelvic floor muscle exercises would have needed more time and that the ball seemed a good method to strengthen the pelvic floor; one mentioned that she would also have been happy with the other group. In the comparison group (seven respondents), five participants stated to have been happy with the allocation and gave as reasons that the exercises had been feasible anytime and anywhere (*n* = 2), that they had been forced to exercise (*n* = 1), and the potential adverse effects of ball use (*n* = 1). The two participants not happy with their allocation stated that they would have liked to get to know the vaginal ball and that ball use seemed easier than performing pelvic floor muscle exercises.

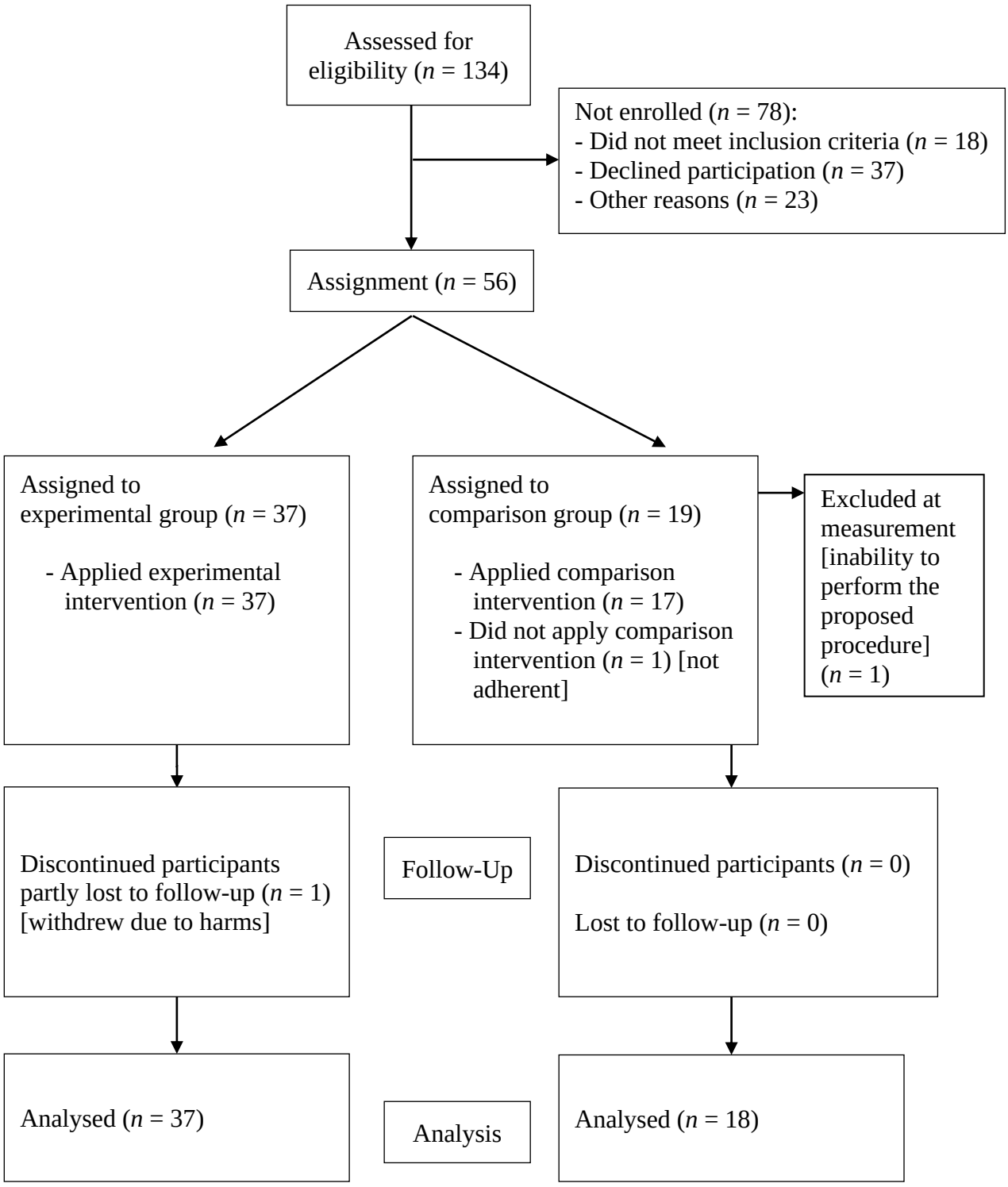
6.4 Completion

Figure 20 summarises the participants' flow through the trial. The participant who had been unable to voluntarily squeeze her pelvic floor muscles at the initial pelvic floor measurement was excluded from the trial with the suggestion to consult her obstetrician and possibly start physiotherapy. One participant in the experimental group withdrew because of adverse events. The last phone contact with this participant was at her medical treatment start when she agreed to be contacted again to check for her wellbeing. After this, she could not be reached any more despite at least 12 attempts to call her and gave the impression that she did not want to be talked to any further.

Considering the 54 trial completers, the mean duration of study participation was 18.6 begun weeks (*SD* 2.0, range 15-25). For completers who were not asked to fill out the online survey, the study ended with the final call approximately two weeks after the final interview or measurement (whichever came last); the study duration for these participants was 15-25 begun weeks (mean 18.7, *SD* 2.5). For completers who were asked to fill out the online

survey and therefore might have got a reminder approximately two weeks after the final call, study duration was potentially longer. Calculated with the last reminder as the end point, the study duration for these participants was 16-23 begun weeks (mean 18.4, *SD* 1.5).

Figure 20 Flow of (potential) participants through the trial (according to Schulz et al., 2010)



6.5 Trial forms and data collection with respect to a future RCT

In this section, the trial forms to be used in a future full RCT are considered first. Then, results for the data collection processes necessary in a future RCT are presented. For the organisation of data collection this is expressed by a description of organisational issues and by the feasibility rates. This is followed by relevant details of selected data collection processes.

6.5.1 Trial forms

Trial forms comprise those necessary for the trial processes and those for data collection of a future RCT. All except three forms underwent refinement during the trial period as this feasibility trial, by its very nature, used nonpiloted forms which revealed their weaknesses during their use. The original forms can be found in the research protocol (Appendix N), the final versions of the forms that were modified and used in this feasibility trial are appended to this thesis. Appendix FF lists all issues with trial forms, gives respective suggestions for a full RCT, and refers to the forms' Appendix numbers.

A form not contained within the trial protocol but created before trial start was the measurement standard for the perineometry appointment (Appendix AA). During the trial, a checklist was designed for the final study meeting (Appendix Y). Administering data collection forms to and collecting them from participants/assessors or using them as PI or assessor mostly went as planned. The administration of the data collection forms is described in detail after the following section on organising appointments with (potential) participants.

6.5.2 Organising appointments with (potential) participants

With (potential) participants, research meetings and measurement appointments at the beginning and end of their participation had to be organised. As CO was flexible, this was easy for the research meetings. Women usually preferred meetings before noon because of childcare availability for their older children. For the information/consent/initial study meeting, all except one woman chose to be visited at home. This one woman was met in public, but at her spontaneous suggestion the meeting finally also took place at her home. Two journeys to visits were in vain as the potential participants were not at home, and two women decided not to take part in the trial at the information/consent visit. The 54 final meetings took place at participants' homes 50 times and four times at a public place.

For the need to bring together an assessor and as many participants as possible at one timepoint, perineometry appointments were more difficult to organise and frequently required back and forth phone calls or e-mails. On one occasion, participant and assessor did not find

each other at the site. To be more assured that women do not miss the appointment and find the assessor, sending participants reminder text messages the day before the measurement appointment was introduced during the course of the trial.

Two feasibility criteria concern the process of data collection. The rate of participants having attended the first pelvic floor muscle measurement within 3 weeks of consenting to take part in the trial is 92.9% (52/56; 95% exact CI [82.7, 98.0]), and Criterion 2 (asking for ≥ 90%) thus is fulfilled. For the then still enrolled 54 participants, the rate of participants having the final data collection within 2 weeks of ending the intervention for Criterion 5 (asking for ≥ 80%) is 66.7% (36/54; 95% CI [54.1, 79.3]). When calculated for data collection within 3 weeks, the rate is 79.6% (43/54, 95% CI [68.9, 90.4]). Criterion 5 thus is not fulfilled. Example reasons for late measurement are participants being away, appointment not offered at suitable time of the day (mornings were preferred), difficulties to organise travel to the site (as moved residence), being back at paid work, or cancellation of appointment; on the side of CO this was being away or having missed to arrange an appointment (once).

6.5.3 Data collection process in detail

For demographic and baseline characteristics, the collection process and tool worked well; in rare cases, it could not be identified whether participants had had a 1st or 2nd degree perineal tear, and the most likely diagnosis from the participant's birth story was then recorded. Administration also worked well for the online survey questionnaires, interview schedules and adverse events calls documentation (although, instead of the three planned calls, as many calls as needed to reach each participant were attempted). The latter was also used to follow participants' progress through the trial by, after each call, placing the form into a folder per date and purpose of next contact. To reach an online survey response rate of 80.8%, 13 (50.0%) of the 26 participants who were sent the survey link needed at least one reminder. Appendix FF lists all issues with data collection processes and gives respective suggestions for a full RCT. Three more noticeable data collection topics are considered in the following.

Pelvic floor questionnaire

When filling out the questionnaire, some of the first participants overlooked items or by mistake ticked "not at all" (the bothersomeness option for women with symptoms) instead of "not applicable–I have no problem" (the option for women without symptoms), or the other way round. This prompted CO to always check the completeness and "correctness" of answers after being handed back the questionnaire.

Question 1 on participant reported pelvic floor muscle strength was tested for its feasibility. Some participants were slightly reluctant to choose a percentage with the argument that they

were not able to express a pelvic floor change in percents but after some encouragement provided an answer; only the participant who was not able to hold the ball in the vagina declined to answer the question at the final meeting. The formulation of the pelvic floor comparison, asking for “now with before this birth”, prompted a discussion about question clarity with two participants who reported a stronger pelvic floor after birth compared to before. For them, the correct question wording would have been to compare the pelvic floor “now with before the recent pregnancy” as pregnancy had brought about the pelvic floor changes.

Adherence

Six participants lost their adherence chart at some point during the trial which then led to replacement forms being completed at least partly in retrospect; also, one participant reported to have filled out the form every two to three days only. Another participant reported to have documented her adherence not immediately but later on the day by which time she “had forgotten” the value she needed to enter; she suggested ticking adherence on the sheet instead of inserting the minutes.

Perineometry

Some issues arose for the perineometric measurements. The assessors fed back that it had sometimes taken quite a while to explain correct contraction to the participant and exercise with her until a clear contraction was palpable. This supported some participants’ initial doubt as to whether they were performing correct contractions during pelvic floor muscle exercising. Two participants were not able to squeeze their pelvic floor muscles at the initial pelvic floor measurement. One was excluded (as detailed in section 6.4), the other assured at the clarifying phone call her ability to squeeze her pelvic floor muscles and blamed the recumbent body position. As her circumstances were conducive in that she only had minor pelvic floor symptoms, she was given a second opportunity which resulted in successful measurements.

In spite of an introduction to the Peritron device for the assessors, measurement errors occurred. For one early participant, the assessor recorded the value of 0 for strength and endurance although a contraction had been palpable. It seems likely that 0 was recorded instead of values under 5, owing to the fact that values under 5 are not saved and shown by the Peritron when looked up after measurement; instead they need to be read off from the display during measurement.

One assessor discarded values for measurements and retested muscle performance when she did not trust the measured value (this happened two times for vaginal resting pressure

and nine times for endurance), detected incorrect technical manoeuvres on the side of the participant (four times), when the probe had slipped out (once), and when one participant wished to repeat measurements after realising a learning effect (once). Another assessor reported repeated measurements for suspect readings four to five times.

The concept of measuring contraction endurance by the number of seconds that the pressure stayed above 5 cm H_2O, with a maximum of 10 seconds, seemingly could not be fully conveyed to the assessors. More than half of all communicated endurance values were over 10, with the highest value being 83 seconds; at the final meeting, the assessors were not able to explain the high values. When probed, one assessor counted the 10 seconds so slowly that they finally were 20 seconds (still not 51, her highest value), and the other assessor remembered a very long contraction (not in line with 10 seconds). It therefore seems as if assessors did not end measurements after 10 seconds. There were also the nine repeat measurements for endurance values not trusted by one of the assessors. When discussing this fact with a second assessor, she agreed and admitted that some of her own values seemed implausible to herself and she questioned whether participants had fully understood the instruction.

Upon discussion of the endeavour to keep them blinded, the assessors reported that participants had revealed their group allocation in two cases. This then prompted one of them to remind participants at the beginning of every meeting not to mention their allocation. Other measurement issues reported by the assessors were accompanying babies/children needing attention and a hurrying participant (once). Technical details reported for the Peritron were that it sometimes had not deleted values after clearing and that a part of the device had broken at the end of data collection. Access to the online perineometry form was not possible as planned (details in data management section 7.1).

Participants‘ ad hoc feedback (not collected systematically) on measurement revealed some further issues:

- Three of them mentioned slight differences between the assessors’ instructions and one (of the three) an insufficient instruction on one occasion. One assessor had developed her own creative instructions to teach correct contractions but had kept to the standard during measurement (clarified at final meeting with assessor).
- For one participant (with dyspareunia), it was difficult to have the vaginal probe inserted on one occasion, and one perineometry measurement was painful for another participant which she ascribed possibly to the birth control pill; one assessor mentioned that it had been difficult to insert the probe for a third participant.

- Two participants described it as difficult to contract the pelvic floor muscles in a lying position, two mentioned difficulties to perform the requested manoeuvres.
- Two participants mentioned that they had been distracted by the baby during the measurement, one of them had the baby lie on her belly. This prompted a discussion with the assessors who confirmed that the babies sometimes had been lying on or near the mothers' chest.
- One participant reported insufficient attention of one assessor to noncontraction of other muscles than those of the pelvic floor. Although assessors counter held the probe against slipping out, one participant mentioned that it had slipped out without the assessor becoming aware of it.

6.6 Interventions

Acceptability and feasibility of the experimental and comparison interventions are considered in the following, with a separate section on intervention uptake and adherence.

6.6.1 Intervention feasibility

As one participant in the experimental group withdrew, the final interview was conducted with 36 participants of this group. Of those, one could not hold the ball inside the vagina as planned and could not contribute to all questions. The following analyses for this group therefore have a sample size of 36 or 35. In the comparison group, one participant was excluded and the final interview was conducted with 18 participants. Of those, one never started exercising and could not contribute to all questions. The following analyses for this group therefore have a sample size of 18 or 17.

Table 11 shows the results for the experiences with both interventions by their easy and difficult aspects named in the interviews and online survey. The written and oral instructions for ball use in the trial seem to have been understandable although not all participants may have read them. A potential (depending on future trial design as delineated in Chapter 10) necessary change in the participant information and consent form is to insert the information that participants are not discouraged from doing their routinely recommended pelvic floor muscle exercises. The written and oral instructions for the pelvic floor muscle training seem to have been understandable, too, but not all participants may have read them. One participant fed back that she would, in retrospect, have liked a more thorough explanation of the requested pelvic floor muscle exercises.

Despite the disadvantages named, 31 of 36 participants in the experimental group (86.1%) could imagine to use[87] the vaginal ball in the future; Box 6 shows their comments. At the final interview, 10 participants in the comparison group were interested in ball use, while four were not interested and four were not sure about this; none of these participants had bought a ball during the trial. All except one participant in the comparison group could imagine continuing the training, e.g. to further strengthen the pelvic floor (*n* = 5) or as a preventive measure (*n* = 3). One participant could only imagine to continue from time to time and not so intensely as she found the training bothersome and felt her pelvic floor does not need training.

[87] The interview schedule question on whether participants would “like to continue” with the intervention was in the conversation changed to whether they could “imagine to continue” as this better met its intention.

Table 11 Comparison of both interventions' feasibility aspects from interview and online survey

	Experimental group	Comparison group
Experience (interview)	24 participants (of 36, 66.7%) gave a positive comment (such as "no problem") and 29 (80.6%) mentioned difficult aspects.	Eight participants (of 18, 44.4%) gave a positive comment about the training (such as "not difficult") or "got accustomed to it quickly" [ID 34]); nine (50.0%) mentioned difficult aspects.
Fit into daily life (interview)	• For 28 participants (of 35, 80.0%), ball use fitted into daily life, for five (14.3%) it did not fit, and 2 (5.7%) were not sure about this. • Comments were: "at the beginning I had to find out when it fit best", "developed a rhythm", "you just build it into your day like medication", "with routine very simple".	• For 11 participants (of 18, 61.1%), pelvic floor muscle training fitted into daily life, for two (11.1%) it did not fit, with five (27.8%) not being sure about this. • One participant found it easier to do the training when on holiday.
Easy aspects (interview)	• Participants found it easiest to use the ball at home (27/35, 77.1%), often during housework, and to a minor extent (5/35, 14.3%) outside of the house. Unusual applications were during dancing or Yoga. • Comments were: "ball was funny and I liked to tell it to my friends", "comfortable and not necessary to take time", "more comfortable than pelvic floor muscle exercises", "for lazy people".	• The 17 participants trained at home (*n* = 14, 82.4%) or in public (*n* = 3, 17.7%) as "nobody sees it", during feeding their baby (*n* = 8, 47.1%), or in the evening in bed (*n* = 5, 29.4%). • Two participants described their process of learning to do correct pelvic floor muscle training.
Difficult aspects (interview)	• 16 of 35 participants (45.7%) considered the constraint to 30 minutes too short a usage period for going out of the house, and named the need to move for 30 minutes at home as sometimes difficult to organise. When leaving home, participants had to think to take the ball with them or where and how to take the ball out; one participant also mentioned that the body positions taken outside (e.g. bending over the child) were conducive to the ball slipping out.	• The answers by 17 participants on difficult aspects were: having to think about the training (*n* = 2) and forgetting to exercise (*n* = 5), child(ren) (*n* = 5), finding time (*n* = 3), willpower (*n* = 3), too many exercises required (*n* = 3), interrupted blocks (*n* = 2), holiday (*n* = 1) and adverse event (*n* = 1) • No disadvantages of the training were named. • One participant stated to have been "far away from

	• To feel bound timewise or to the house was therefore considered a disadvantage by seven of 35 participants (20.0%). Eight participants (22.9%) said that ball use was less feasible during holiday, at the weekends, or during times of stress. Seven (20.0%) mentioned the need to think about ball use as disadvantage or difficulty, five their child(ren). Two participants reported that they had walked around only for ball use or had even got up from bed at night to fulfil the intervention prescription.	increasing" the dose (ID 27).
Online survey Opinion on pelvic floor muscle strengthening by intervention after 12 weeks	14/18 questionnaires completed, 13/14 participants answered the question: • Five called the intervention easy or practical. • Seven gave other positive comments such as "ball use is fun" or "can be recommended". • One participant questioned the lasting effect of the changes and thinks that ball use needs to be continued.	7/8 questionnaires completed, 6/7 participants answered the question: • All comments were positive: "it works for sure", "I find strengthening the pelvic floor important", "good idea", "it for sure was helpful to me", "can do them always and everywhere", "feasible without much effort", "I have learned to use small breaks and am exercising regularly". • Two participants critically said that "one always has to motivate oneself" and that "it often was difficult to do the exercises three times a day with the everyday stress with a baby".

Box 6 "Can you imagine to use the vaginal ball in the future?" (*N* = 36, multiple answers possible)

Yes:

- Because it is practical (*n* = 5)
- To further strengthen the pelvic floor (*n* = 4)
- Because it helped (*n* = 4)
- If it helps (*n* = 3)
- To enhance sexuality (*n* = 2)
- If the pelvic floor gets worse (*n* = 2)
- With heavier balls (*n* = 2)
- As a preventive measure (*n* = 1)
- Other (*n* = 6): "because of help for haemorrhoids (after my first baby they were not good for three years)", "interesting experience–not off-putting", "I would if I was fitter–it does no harm", "still tempting–a good possibility if it stayed inside", "because I like to do something for the pelvic floor, it is comfortable to wear; if it does not help it does no harm", "why not"

No:

- No need for training (*n* = 1)
- Only if pelvic floor gets worse–because no detectable effect at the moment (*n* = 1)

Not sure: (*n* = 3)

- At times when the pelvic floor gets worse (*n* = 2)
- Only if it worked objectively (*n* = 1)
- Rather not for fun; would prefer sports and pelvic floor muscle training as this can be done everywhere (*n* = 1)

Specific issues in experimental group

Insecurity, at least initially, about the height at which the ball needed to be placed in the vagina was mentioned by five participants. Losing the inserted ball at some point in the trial was mentioned by nine. This happened the first times of ball use or was connected to the use of lubricant, day of the menstrual cycle (with more discharge), or strain on the pelvic floor. On the adherence sheets, five participants indicated that they had not been able to hold the ball as long as planned. This happened between one and 11 times per participant (mode 1, median 2); the two reported reasons were that the ball had flopped out on the toilet and that the pelvic floor was weaker in the evening. Having forgtten to take the ball out was mentioned by five (of 36, 13.9%) participants; five participants had set a timer to be reminded to take the ball out.

Two participants' balls had fallen into the toilet; while one participant asked for a replacement ball, the other had cleaned the ball and used it again, telling about this only at the final interview. One participant forgot to take the ball with her when travelling, and one lost her ball and needed a replacement ball.

Handling the ball was considered easy by all but one of 36 participants, but four mentioned issues with cleaning. The cleaning issue was confirmed by other participants when probing for ideas to enhance the ball: Altogether, the ball's dimples[88] as a feature that could be enhanced for hygiene reasons was named six times, and three and two participants respectively called the retraction cord and the ring around the ball[89] impractical with respect to cleaning. One participant suggested a different colour than white for the cord as it had become greyish during use, and a longer retraction cord was suggested by three participants. One participant suggested to make the ball smaller and one to make it softer.

An issue emerging during the trial was ball handling after episodes of vulvovaginal symptoms (see section 8.3). Upon the first case, hygiene staff of the AKH Vienna was contacted for advice on ball disinfection. However, the advice given is unfeasible in a trial as participants usually presumably would neither have an appropriate disinfectant at home nor the knowledge about correct disinfection. "Lay disinfection" with hygiene spray or similar equipment (as mentioned in Intimina (no date) or MAPA GmbH (no date-c)) is not sufficient, and boiling the ball (an effective disinfection method feasible at home) not possible (Intimina, 2015, no date). Except for one who insisted on boiling the ball, trial participants therefore were provided with a new ball.

Thirteen participants (of 36, 36.1%) performed pelvic floor muscle exercises additionally to using the ball; six of them (16.7% of 36) regularly (from three times/week to daily), five at irregular intervals (from very rarely to one to two times a week), and two only at trial start. The nature of the regular exercises was not enquired in depth but likely comprised the gentle exercises routinely recommended in Austria.

6.6.2 Intervention uptake and adherence

Subtracting the participant excluded after her failed initial pelvic floor measurement, the corrected population of interest for Feasibility criterion 3 on intervention start is 55 participants. The rate of participants having started with the intervention within 4 days of the initial pelvic floor muscle measurement is 81.8% (45/55, 95% CI [71.6, 92.0]). Within 7 days, 89.1% (49/55, 95% CI [80.9, 97.3]) of the participants had started with the intervention. Criterion 3, requiring a rate of ≥ 90%, thus is not fulfilled. Reasons that could be identified as to why participants had not started with the intervention sooner were being on holiday or frequently away, family obligations, the intention to start on a Monday, and the baby needing intensive care.

[88] These are two opposite concave ball indentations of which one serves to fix the retraction cord to the plastic ring around the ball and the other to fix the retraction cord of another ball in case two balls are combined, see Figure 13.
[89] This is a plastic ring around the ball to secure the retraction cord to the ball, see Figure 13.

After intervention start, one participant from the experimental group withdrew, and the participant who was unable to hold the ball inside the vagina for more than 3-15 minutes was excluded from the adherence analysis. Therefore, adequate adherence to the interventions, calculated according to the planned analysis as adherent on at least 80% of the days during the intervention period, was 47.2% (25/53, 95% CI [33.8, 60.6]) overall, 51.4% (18/35) in the experimental and 38.9% (7/18) in the comparison group.

The adherence sensitivity analysis with the newly developed adequate adherence definition resulted in an adherence rate of 60.4% (32/53, 95% CI [47.2, 73.6]) overall, 62.9% (22/35, 95% CI [46.9, 78.9]) in the experimental and 55.6% (10/18, 95% CI [32.7, 78.6]) in the comparison group. Feasibility criterion 4b (≥ 80% adhere adequately) thus was not fulfilled. In contrast, Feasibility criterion 4a (≥ 80% keep to assigned intervention group) was fulfilled as all participants (95% CI [93.4, 100]) kept to the assigned group. Reasons for nonadherence named by participants are listed in Box 7.

Box 7 Reasons given by participants for nonadherence (multiple answers possible)

Experimental group (*n* = 18):

- Forgetting (*n* = 5)
- Holiday (*n* = 5)
- Illness (own or family member, *n* = 5)
- Vulvovaginal symptoms (*n* = 3)
- Being out of the house all day (*n* = 3)
- Not finding 30 minutes (*n* = 2)
- Ten other reasons were: “not used with muscle soreness”, “ball slipped out”, “on the weekends everybody was at home”, “house moving”, “did not dare to go over 15 minutes because of pressure symptoms”, “motivation had gone after holiday, also because it did not seem to help”, “motivation decreased when being back at paid work”, “forgotten during daytime and no desire to use it any more in the evening”, “had planned to use it during housework in the evening when children were asleep, but when I could not sense the vibrations I wanted to use it when out for a walk, but then was not able to put it in before leaving the house”, and “fun had decreased towards the end of the trial period; 10 weeks would have been easier, 12 is a bit long”.

Comparison group (*n* = 9):

- Forgetting (*n* = 6)
- Interrupted blocks (*n* = 2)
- Holiday (*n* = 1)
- Two other reasons were: “sometimes remembered too late on the day or then not in the mood for it any more” and “one´s weaker self”.
- The participant who had never started exercising reported that, when asked by her obstetrician, she had not been a 100 percent convinced but participated for his sake and had imagined participation easier.

The ball was sometimes used longer than the intended 15 and 30 minutes respectively by 27 participants (calculated of 35 available adherence sheets). “Longer” was either not specified or went from 31 minutes over 40-60 minutes up to 7 hours. If reasons were provided, these were: was out of the house and could not take the ball out, no time to take the ball out as caring for child, forgotten to take ball out, and sitting time included. One participant had usually left the ball for two to three hours but had noted 30 minutes as “moving time“. Some participants noted the exact number of minutes, but those who always noted 30 minutes may have kept the ball in for longer or shorter. In the comparison group, four women (of 16 available adherence sheets) sometimes performed more than the planned number of exercises. The two reasons given were to have understood that four exercise blocks daily were requested and to catch up for fewer exercises on other days.

Summary

Preselected women’s interest in the study was high, and an overall recruitment rate of 48.3% (95% CI [39.2, 57.4]) of eligible women at the researcher level fulfilled Feasibility criterion 1. Recruitment success and failure at each recruitment level could be identified. The selection criteria proved feasible and only need few amendments.

The sample was mainly Caucasian, highly educated and socially well situated. Urinary incontinence symptoms were reported by 21 of 56 participants. Almost all participants had experience with pelvic floor muscle exercises, and most had received information about exercising at the reference birth. A third of the participants performed pelvic floor muscle exercises at least two to three times a week at trial start. About two thirds of the participants had heard of vibrating vaginal balls and seven had already used one.

At trial start, participants mostly had a high opinion of pelvic floor muscle exercises and a positive attitude towards the balls. Their main motivation for participation was to do something for the pelvic floor, an expected knowledge gain, and being intrigued by the topic. They had a high preference for the experimental group and were highly motivated to perform their intervention during the trial period. Randomisation worked smoothly. One participant was excluded when found to be ineligible after inclusion, and one withdrew, resulting in 54 of 56 participants completing the trial.

This feasibility trial suggests that appropriate measures, measurement instruments and data collection methods were used. Refined versions of the data collection forms could be designed. The most challenging part of data collection proved organising perineometry appointments. Nevertheless, Feasibility criterion 2–initial perineometry measurement within 3 weeks of consenting to take part in the trial–could be fulfilled; Feasibility criterion 5, asking

for the final data collection to be completed within 2 weeks of ending the intervention, could not be fulfilled.

The trial interventions generally proved feasible. Feasibility criterion 3 (intervention start within 4 days) was nearly fulfilled. Disadvantages named for the use of vibrating vaginal balls were the need to think about it and the restriction to 30 minutes. Routine ball handling was reported as easy. Pelvic floor muscle training was rated positively, but forgetting to do it was a barrier to adherence. Feasibility criterion 4b on adherence was not fulfilled, even not with the application of newly developed wider margins of error for its definition. On the other hand, the ball was sometimes used longer than the intended time span, and more exercises than intended had sometimes been performed.

The following results chapter covers trial management and resources findings.

7 RESULTS II – TRIAL MANAGEMENT AND RESOURCES

Following the framework of Thabane et al. (2010), this second results chapter presents the findings on trial management and resources. By looking at these particular trial feasibility aspects, it contributes to the overall research question on how best to prepare and perform the intended future RCT.

7.1 Trial management

Identified human, organisational and data management issues and difficulties are described in the following.

Human management

Human management in this feasibility trial was needed for trial administration, recruitment sites/professionals, (potential) participants, research visits and pelvic floor assessors. As human management with regard to recruitment sites/professionals and (potential) participants was considered within the trial process findings in the previous two chapters, human management in this section would only concern administrative and data collection staff. Due to the PhD nature of the project, trial administration and research visits were done by CO. As her self-management is out of the range of this trial's analysis, human management results here finally only concern the pelvic floor assessors.

First, one or more pelvic floor assessors had to be found[90]**.** The first nine measurements were done by EH. As planned, he then organised a final year medical student (SC) with a strong interest in a gynaecological/obstetrical career to help with the assessments. Additionally, midwife EB was found via the AKH Vienna midwifery team. No funding was available to pay the assessors; their motivation to support the project was the promotion of clinical research, the learning opportunity, and a contribution to their professional career. All three assessors were insured for the research activities, got a theoretical and practical introduction to the measurement device and standard, and emergency procedures were clarified with them. Risk assessment according to City University London policy was performed for CO (PI) but not for the assessors for whom this should also have been undertaken (no negative consequences evolved).

For the measurements, a rota had to be organised, taking into account the assessors' full time working schedule and participants' time needs, with the aim to gather as many participants per date as possible. This was a challenge because of illness of assessors,

[90] Inability to find outcome assessors was the second reason why the trial could not be run in the UK, apart from recruitment failure at the management level (see page 123).

participants or their family members, and time constraints by work duties, holidays, healthcare appointments, lack of childcare availability etc. In 41 sessions, 67 measurements were performed by SC, 35 by EB and nine by EH. Although the perineometry data analysis and interpretation were performed by CO, the assessors were consulted to contribute their experience, knowledge and ideas.

Organisational management

Organisational management issues were identified with respect to ethics, trial venues, and material needs.

Ethics

The ethics committees in Austria and the UK have different guidelines, procedures and a different cultural background. Two delicate ethical issues were the Peritron only having CE marking[91] for the vaginal probe and vaginal measurement as an intimate topic; both were more scrutinised in the UK but could be resolved to the satisfaction of both committees.

During the study period, two ethics amendments were needed. The first one concerned introducing SC and EB as external assessors for perineometry after trial start. The second amendment concerned accepting women as participants who declined the six week postpartum check when one woman interested in participation had declined this check because she felt in good health. The amendments corresponded to the applicable ethical guidelines and were clarified with both ethics committees.

The trial was conducted in accordance with the research protocol, with two exceptions. As is explained in detail in the data management section below, one exception was a failure in keeping perineometry data anonymous outside of the electronic data collection system (no known negative consequences). The second change was that, contrary to the original plan to only interview the first half of the participants, shortened interviews before and after the intervention were also performed with the second half of the participants. As this only meant a burden of approximately 15 more minutes at the final visit for each respective participant, it was considered minor and ethics amendment was not sought. Due to the feasibility nature of the trial, minor changes were made in the trial forms' wording as documented in Appendix FF.

[91] CE marking signifies that products sold in the European Economic Area "have been assessed to meet high safety, health, and environmental protection requirements" (European Commission, 2018, https://ec.europa.eu/growth/single-market/ce-marking_en). "CE" stands for "Conformité Européenne" in French which means "European Conformity" (Wellkang® Tech Consulting, 2017).

Venues

The research office was the home of CO and her office at City University London. For the information/consent/initial and final study meetings, participants were offered home visits or to meet at a public space. They almost all preferred their homes, only four meetings took place at public places (see section 6.5). For perineometry measurements, two rooms were available at the AKH Vienna: One was a ward bathroom with a noninspiring atmosphere–in the words of a participant an “ugly room, but lockable” (no ID as colloquial); the other room was a cosy antenatal care room. This room was only available in the afternoon; the preferred time to use the bathroom was in the afternoon as well because the ward was busy in the mornings and one assessor did not feel welcomed during these hours.

Materials

At the start of this feasibility trial, the Peritron was not sold in EU countries as only its probe had CE marking; it thus was imported via a non-EU country where it could be purchased. In the meantime, it has been fully CE certified and is delivered to EU countries, and its purchase in a full trial should not be a problem. As secure storage of the Peritron posed a challenge in this feasibiliy trial, storage of the Peritron (and other material needed for measurement) should be considered before the start of a full trial.

Refunding transport tickets to participants at the perineometry measurement did for two reasons not run as planned. One reason was that the tick box “Transport tickets refunded” in the online data collection form could not remind assessors as they could not access the forms in the measurement rooms, and without this reminder it seemed easy for them to forget the tickets. The other reason lay in the logistical circumstances of CO doing this research as a PhD. Tickets therefore were in many cases refunded to participants at the final study meeting. This was an acceptable solution for the feasibility trial but needs to be thought through for a full RCT.

Data management

An unexpected data processing situation encountered was that neither computers nor Internet access were available in the measurement rooms to enter the data into the electronic perineometry form. The assessors therefore noted the results on paper and entered them into the electronic form later at another site. As they did not know the participants’ IDs (nor did the participants themselves), they noted their names and date of birth together with the results, meaning that these data were not kept anonymous (but confidential) outside of the electronic system. After the end of data collection, the assessors were reminded to destroy any data notes left.

As interview notes were transcribed as soon as possible from the paper forms into the electronic database, content remembered but not jotted down during the interviews was nevertheless entered electronically. This led to entries on paper forms and in the electronic database not being fully congruent and the electronic form being more comprehensive. As CO was not experienced in such matters, no formal system of data files naming and dating was planned or used, neither was there coherent version numbering of documents (except for the participant form as explicitely required by the ethics committee of the Medical University of Vienna). Quantitative data analysis via SPSS and R worked well, some codes needed relabelling due to CO's lack of coding experience. Qualitative data analysis via Excel went smoothly, and there were no problems with the other computer programmes used.

7.2 Trial resources

Resources needed for this feasibility trial are summarised in Table 12. A few resources issues were encountered unexpectedly and might be of importance in planning a full RCT:

- Originally it was hoped that the ball manufacturer Intimina would fund the balls and lubricant. Upon information about the trial plans however, the contact person of the company communicated that Intimina would prefer the balls to be used as described in their instruction materials, which is to have the ball inserted during performing contractions (Intimina, 2018c). As this form of use was not the underlying research interest in this project, this was not done and Intimina did not contribute to funding.

- Four replacement balls had to be purchased and were sent to the participants by mail: two to replace the ball after vulvovaginal symptoms, one because the participant had misplaced the ball and could not find it any more, and one because the ball had fallen into the toilet. From the three remaining participants with vulvovaginal symptoms who would also have needed a replacement ball, one withdrew, one purchased the replacement ball herself, and one insisted on disinfecting the ball by boiling it.

- As the first final interviews showed that the initially provided lubricant package of 200 ml per participant was not needed, another lubricant brand in a 150 ml package was purchased. This amount was not used up by participants either but there was no economically priced smaller lubricant package available.

- Some participants had an annual public transport pass and would not have needed transport refund as they did not incur costs. However, as the formulation in the participant information and consent form was that participants would incur costs and be handed two tickets per attendance, it was refunded to all participants.

Table 12 Resources needed for feasibility trial with 56 participants

Category	Item & amount (if applicable)
Facilities	- Office with computer, printer and web accessibility - Room for pelvic floor measurements (lockable) - Waiting area with toilet - Lockable storage place for Peritron - Library access
Staff (+ staff time)	- Administrative: organising study visits and measurement appointments - Scientific: for recruitment, data collection (duration of initial study visit 1-1.5 hours, of final study visit 0.5-1 hour, phone calls, mails), data processing and analysis, report writing - Blinded assessors: 10-15 minutes per measurement - Randomisation support - Statistical support - Liability insurance for hands-on scientific staff
Materials	<u>For administration and data collection/analysis:</u> - Mobile phone - Computer programmes: Office 365, SPSS, R, endnote - Subscription to web survey application software - 328 recruitment sheets (1 colour page A4) to distribute - 23 (2 pages, 1 colour) recruitment forms for health professionals - 581 information/consent forms (11 pages, cover page in colour) - 9 paper data collection forms (23/21 pages for long/short interviews respectively, 3 in colour) per participant - 30 pages recruitment phone call documentation sheets for researcher/participant list - 1 pelvic floor picture (colour) - Stationeries (envelopes, folders, writing material) - Postage for sending 4 replacement balls and a transport ticket by mail - Catering bills for meetings with recruitment professionals/ assessors/participants <u>For experimental intervention per participant:</u> - 1 vaginal ball - 1 lubricant (200 or 150 ml respectively) - 1 antibacterial soap (300 ml) + 4 replacement balls for 37 participants in experimental group <u>For measurements:</u> - 1 Peritron <u>Per participant:</u> - 4 examination gloves (some latex free in stock) - 2 latex (ultrasound) probe condom covers (some latex free in stock) - 4 (2x2) ml lubricant - 2x exam table paper - 2x2 disinfection wipes for Peritron and examination table - Spares of each item

Category	Item & amount (if applicable)
Transport	Per participant: - 4 public transport tickets to refund measurement attendance - 4 public transport tickets for home study visits (except car for two difficult to reach destinations) 2% spares for wasted and re-visits

Summary

Management findings comprise human, organisational and data management. The most relevant organisational findings for a future RCT are that vaginal manometry as an intimate measurement is considered differently in different countries and that perineometry data protection must be assured. Trial resources findings (anticipated and unexpected costs) could be compiled and can serve as the basis for budget calculation for a full RCT. The following results chapter deals with the clinical findings.

8 RESULTS III – CLINICAL FINDINGS

The clinical findings in this third results chapter (still following the framework of Thabane et al. (2010)) first cover the participants' clinical experience with the experimental intervention and by this address the research question on the participants' perspectives on and experiences with the interventions. After this, data from PROMs, perineometry and adverse events screening contribute to the preliminary results regarding pelvic floor muscle performance and harms. This answers the research question on harms associated with the experimental intervention. It also provides essential information for a full RCT sample size calculation, thus contributing to the research question on the number of participants needed for a full trial. With all these findings, this third results chapter contributes to the overall research question on how best to prepare and perform a full RCT.

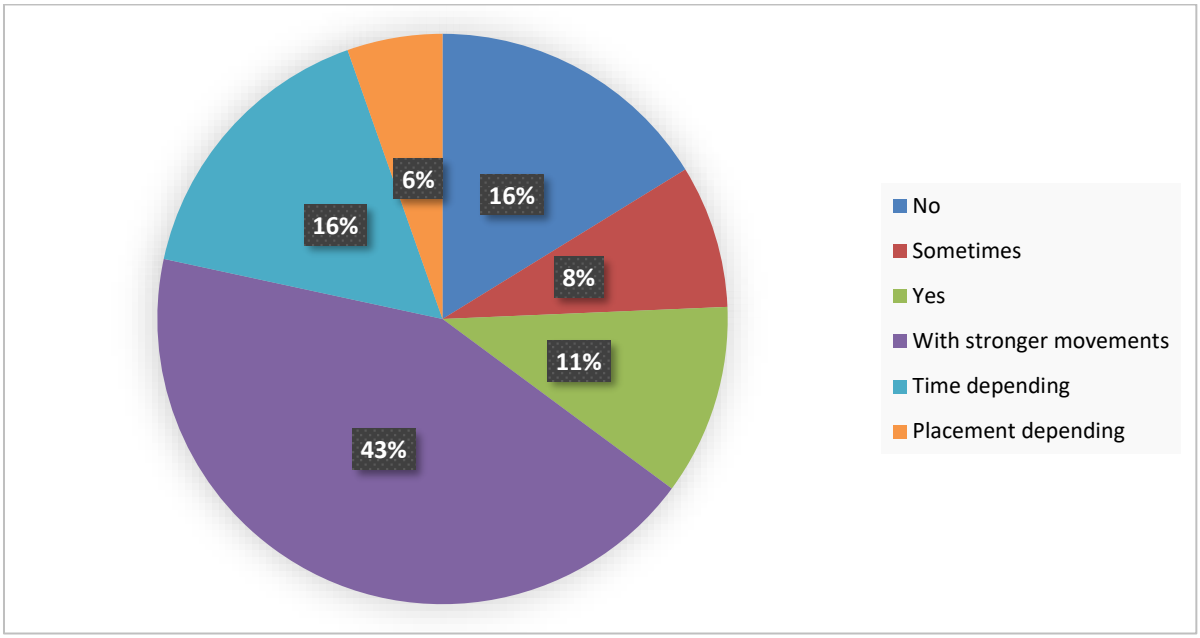
8.1 Participants' clinical experience with vibrating vaginal balls

As one participant in the experimental group withdrew, the final interview was conducted with 36 participants in 18 long and 18 short final interviews. Ten of these participants (27.8%) stated that they did not or barely notice the inserted ball, whereas four (11.1%) felt a downwards pressure from the device. For two (5.6%), noticing the ball depended on the (time of the) day, reporting that it was more effort to hold the ball after preceding physical activity or on the days of the menstrual cycle with more discharge, or in the evening; for one of the two, ball use was less or not possible in the evening because of then losing the ball. One participant could never passively hold the ball inside for more than 3-15 minutes at all, and active holding was not possible to her for longer than one minute.

Of 36 participants, 20 (55.6%) did not notice a pelvic floor reaction during ball use. Four (11.1%) confirmed a sensation of pelvic floor activity; one of them described this activity like at ovulation or shortly before the period, during and for several hours after ball use; another one "like in other body muscles after going for a walk" after ball use (ID 21). Twenty-five participants (69.4%) did not notice any pelvic floor contraction, but four (11.1%) described feeling an enduring contraction. Voluntary contractions to keep the ball inside were performed by eight participants (22.2%), with five (13.9%) performing repeated contractions. Five of 36 (13.9%) felt the sensation of the ball wanting to slip out of the vagina, four (11.1%) felt it sometimes, whereas 19 (52.8%) did not feel it. Commentaries on this sensation specified this as: only the first times of ball use (n = 7), only with increased intraabdominal pressure (n = 6), if the ball was put in too low (n = 6), and only with lubricant (n = 4). Nine participants (25.0%) actually had the ball slip out at some point during use although this mostly was an infrequent occurrence and in three cases was an initial issue only (see section 6.6.1).

Not all participants, namely 30 of 36 (83.3%), felt the ball's vibrations. This could be further specified to a general yes (*n* = 4), sometimes (*n* = 3), only with stronger movements (*n* = 16), timewise at the beginning of the 12 weeks or 30 minutes (*n* = 6), and according to the positioning of ball (*n* = 2) (Figure 21).

Figure 21 "Did you feel the vibrations?" (*N* = 36) (multiple answers possible)



8.2 Pelvic floor muscle performance

After a look at participant baseline characteristics by group, pelvic floor muscle performance is shown for participant reported and perineometry outcomes. All results are presented in descriptive form for available cases, with effect sizes given for the most relevant ones. Unless indicated as PP analysis, the analyses are mITT analyses. The analysis for pelvic floor muscle performance in this trial comprises a sample of 55 participants–37 in the experimental and 18 in the comparison group. Because of a few missing values (detailed in Appendix BB), sample size of the calculations slightly varies. The reader is reminded that the feasibility nature of this trial means that the performance results are exploratory and preliminary. Further, when interpreting the clinical results, the uneven 2:1 sample distribution must be taken into consideration.

8.2.1 Participant baseline characteristics by group

The descriptive group comparison in Table 13 presents demographic and clinical participant characteristics at trial start. This shows that the intervention groups were mostly similar for

the variables measured at baseline. Only two variables differed: the rate of 2nd degree tears at the reference birth with 11.4% (4/37, experimental group) vs 0% (comparison group), and the UI rate with 29.7% (11/37) in the experimental vs 50.0% (9/18) in the comparison group.

Table 13 Demographic and clinical participant characteristics per group at trial start

Characteristic		**Experimental group**[a] (*n* = 37)	**Comparison group**[b] (*n* = 18)
Highest completed education *n* (%)	University degree High school Vocational middle school or apprenticeship	28 (75.7) 6 (16.2) 3 (8.1)	13 (72.2) 3 (16.7) 2 (11.1)
Occupational status *n* (%)	Paid work (maternity leave) Housewife Student	32 (86.5) 5 (13.5) 0 (0.0)	15 (83.3) 2 (11.1) 1 (5.6)
Age (years)	Mean (*SD*) Range	33.7 (4.7) 21-41	32.5 (3.9) 26-41
BMI[c]	Mean (*SD*) Range	22.1 (2.1) 17.9-27.8	23.1 (3.5) 18.3-32.8
Parity *n* (%)	I II III/IV	17 (45.9) 18 (48.6) 2 (5.4)	7 (38.9) 10 (55.6) 1 (5.6)
Completed weeks post partum at inclusion	Mean (*SD*) Median Range	12.7 (5.1) 11 6-22	12.2 (4.8) 10 7-24
Mode of birth *n* (%)	Spontaneous Ventouse	35 (94.6) 2 (5.4)	16 (88.9) 2 (11.1)
Birth injury *n* (%)	Episiotomy Tear 2nd degree None	2 (5.4) 5 (13.5) 9 (24.3)	1 (5.6) 0 (0.0) 5 (27.8)
Birth weight newborn (g) *n* (%)	Mean (*SD*) Range > 4500	3412 (425) 2620-4530 1 (2.7)	3399 (343) 2800-4160 0 (0.0)
Past mode of birth *n* (%)	Ever ventouse Ever caesarean section	1 (2.7) 4 (10.8)	0 (0.0) 1 (5.6)
Past birth injury *n* (%)	Ever episiotomy Ever 2nd degree tear Ever 3rd degree tear	1 (2.7) 4 (10.8) 0	2 (11.1) 2 (11.1) 0

Characteristic	Experimental group[a] (*n* = 37)	Comparison group[b] (*n* = 18)
Weight newborn ever > 4500g *n* (%)	1 (2.7)	0 (0.0)
Breastfeeding *n* (%)	36 (97.3)	17 (94.4)
Cigarette smoking *n* (%)	1 (2.7)	1 (5.6)
Any previous urinary incontinence *n* (%)	20 (54.1)	11 (61.1)
Urinary incontinence *n* (%)	11 (29.7)	9 (50.0)

Note. SD = standard deviation; min = minimum, max = maximum.
[a]One participant equalling 2.7%.
[b]One participant equalling 5.6%.
[c]Normal weight BMI range 18.50-24.99 (World Health Organization, 2018a).

8.2.2 Participant reported outcomes

The PRO's reporting starts with Question 1 on participant reported pelvic floor muscle strength and postintervention Question 2 on enhanced pelvic floor performance. This is followed by a summary of the results of participants' answers to the four symptom and their bothersomeness questions. The section ends with the results of the ICIQ-UI SF and the qualitative PRO results from the interview.

Pelvic floor muscle strength (Question 1)

Table 14 shows participant rated pelvic floor muscle strength before and after the intervention. At trial start, the pelvic floor was considered stronger than before birth by three (of 55) participants, and at trial end by four (of 53; one was the same participant at start and end, the other ratings came from different participants). These participants therefore rated pelvic floor muscle strength higher than 100%–the strength of the identical participant rose, that of the other two participants at trial start diminished, that of the three at trial end rose). Within all participants, 36 (83.0%) of the 53 measurements rose and nine (17.0%) stayed identical. Eight values (15.1%) dropped between 5% and 40% (absolute values, relatively between 7.7% and 36.4%). Of the eight reduced values, six were in the experimental and two in the comparison group; of the nine identical values, seven were in the experimental and two in the comparison group.

In the experimental group, the mean value rose from 76.5 to 85.4, and in the comparison group from 72.8 to 86.9. The change scores predicted from the statistical model are 9.0 (*SD*

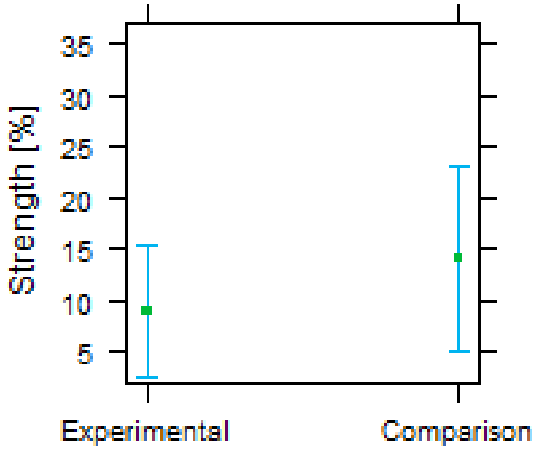
15.9) and 14.1 (*SD* 24.0) respectively. The change score difference of -5.1, 95% CI [-18.0, 7.8], delineates a smaller rise in the experimental group. Figure 22 presents the change scores and their 95% CIs graphically.

Fot this potential future primary clinical outcome, a PP analysis was performed to see whether there was a possible effect according to the reported adherence to the intervention protocol. This PP analysis included 22 participants in the experimental and 10 in the comparison group; the change scores predicted from the statistical PP model are 11.4 (*SD* 17.7) and 23.4 (*SD* 21.8) respectively. With a change score difference of -12.0, 95% CI [-28.8, 4.8], the PP analysis delineates a directionwise identical but larger effect size than the ITT analysis, thus supporting the mITT analysis.

Table 14 Participant rated pelvic floor muscle strength (expressed as percentage compared to before birth) before and after the intervention

	Before intervention Mean (*SD*) Range	**After intervention** Mean (*SD*) Range
Experimental group (*n* before = 37, *n* after = 35)	76.5 (18.9) 25-110	85.4 (16.0) 60-140
Comparison group (*n* = 18)	72.8 (20.2) 50-110	86.9 (16.5) 65-130

Figure 22 Change scores of participant rated pelvic floor muscle strength per intervention group, mean with 95% CIs (*n* = 35 experimental/18 comparison)

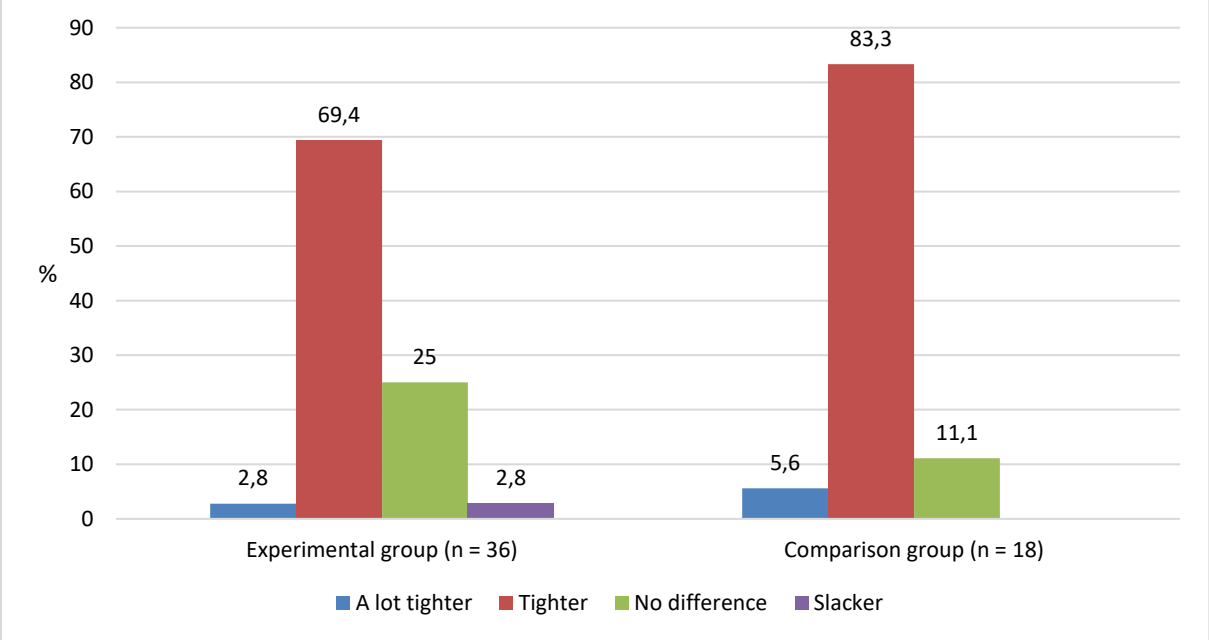


Pelvic floor muscle improvement (postintervention Question 2)

Figure 23 shows the results on pelvic floor muscle improvement. Thereby, no participant chose the category “a lot slacker“. The “tighter” choice was chosen more frequently in the comparison group. “No difference” was found more often in the experimental group, in which also the only “slacker” was chosen.

To calculate the RR, the four categories were dichotomised into the categories “tighter” (including “tighter”/”a lot tighter”) versus ”not tighter” (including “same”/”slacker”). This resulted in 26 “tighter” and 10 “not tighter” in the experimental and 16 “tighter” and two “not tighter” in the comparison group. The risk difference (of not reporting a tighter pelvic floor) is 16.7% (27.8% [experimental] vs 11.1% [comparison]), meaning that participants in the comparison group were 1.2 times (RR 1.23 [0.95, 1.60]) more likely to report increased strength as compared to the experimental group.

Figure 23 Participants’ rating of pelvic floor muscle change after the intervention period (*N* = 54)



Symptom and bothersomeness questions

The answers to the symptom questions were dichotomised into the categories “ever experienced” (including “sometimes”/”frequently”/”always”) versus ”never experienced”, those to the bothersomeness questions into “at all” (including “a little”/”quite”/”very”) versus “not at

all". Table 15 shows the ensuing descriptive results, and the symptom values are also presented graphically in Figure 24.

Although the comparison group shows higher initial values in three of the four symptom questions, the group comparison revealed a greater reduction in this group (cave: few participants). The only symptom to have a frequency rise was "vaginal pressure" in the experimental group. In Table 15, it can be seen that vaginal looseness/laxity and pressure did bother more after the intervention in the experimental group, with wind bother decreasing, whereas wind bother increased in the comparison group while the other two decreased (loss of stool was only experienced by one participant and thus cannot be considered for this outcome). A raw bothersomeness scores comparison (not shown) revealed that the highest category ("very") was unused at the final measurement whereas the "not at all" category diminished for all outcomes.

Figure 24 Descriptive results for the symptoms answer "ever experienced"

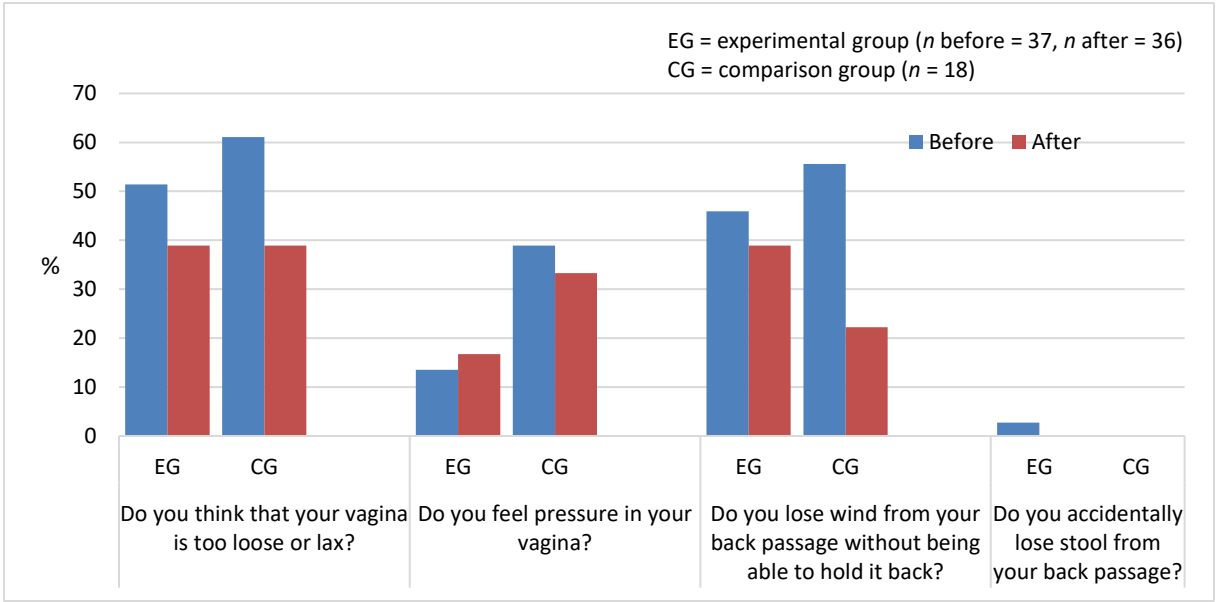


Table 15 Dichotomised descriptive results for symptom and bothersomeness questions

Question		Experimental group		Comparison group	
		Prevalence Symptom: ever experienced Bother: at all			
		n/n[a]	%	*n/n*[a]	%
Do you think that your vagina is too loose or lax?	Pre Post	19/37 14/36	51.4 38.9	11/18 7/18	61.1 38.9
If yes: How much does this problem bother you?	Pre Post	17/19 14/14	89.5 100.0	10/11 6/7	90.9 85.7
Do you feel pressure in your vagina?	Pre Post	5/37 6/36	13.5 16.7	7/18 6/18	38.9 33.3
If yes: How much does this problem bother you?	Pre Post	3/5 6/6	60.0 100.0	5/7 4/6	71.4 66.7
Do you lose wind from your back passage without being able to hold it back?	Pre Post	17/37 14/36	45.9 38.9	10/18 4/18	55.6 22.2
If yes: How much does this problem bother you?	Pre Post	12/17 8/14	70.6 57.1	5/10 4/4	50.0 100.0
Do you accidentally lose stool from your back passage?	Pre Post	1/37 0/36	2.7 0.0	0/18 0/18	0.0 0.0
If yes: How much does this problem bother you?	Pre Post	1/1 n.a.	100.0 n.a.	n.a. n.a.	n.a. n.a.

Note: Pre = preintervention, Post = postintervention.
[a]For the symptom questions, the denominator is the intervention group, for the bothersomeness questions it is the number of participants having answered “yes” at the symptom question.

ICIQ-UI SF results

Table 16 shows the ICIQ-UI SF sum score and the urinary incontinence rate before and after the interventions per group. The score did not change in 35 participants, diminished in 17, and two participants in the experimental group initially had an ICIQ-UI SF sum score of 0 and at trial end of 3. The sum score mean reduction was larger in the comparison group than in the experimental group (1.5 vs 0.8), as was the decrease of the urinary incontinence rate (11.1% vs 7.5%). On the ICIQ-UI SF bother scale, participants overall scored a mean of 1.1 (*SD* 2.1, range 0-9) at trial start and of 0.6 (*SD* 1.6, range 0-8) at trial end. Figure 25 shows the qualitative ICIQ-UI SF results before and after the interventions for all participants per

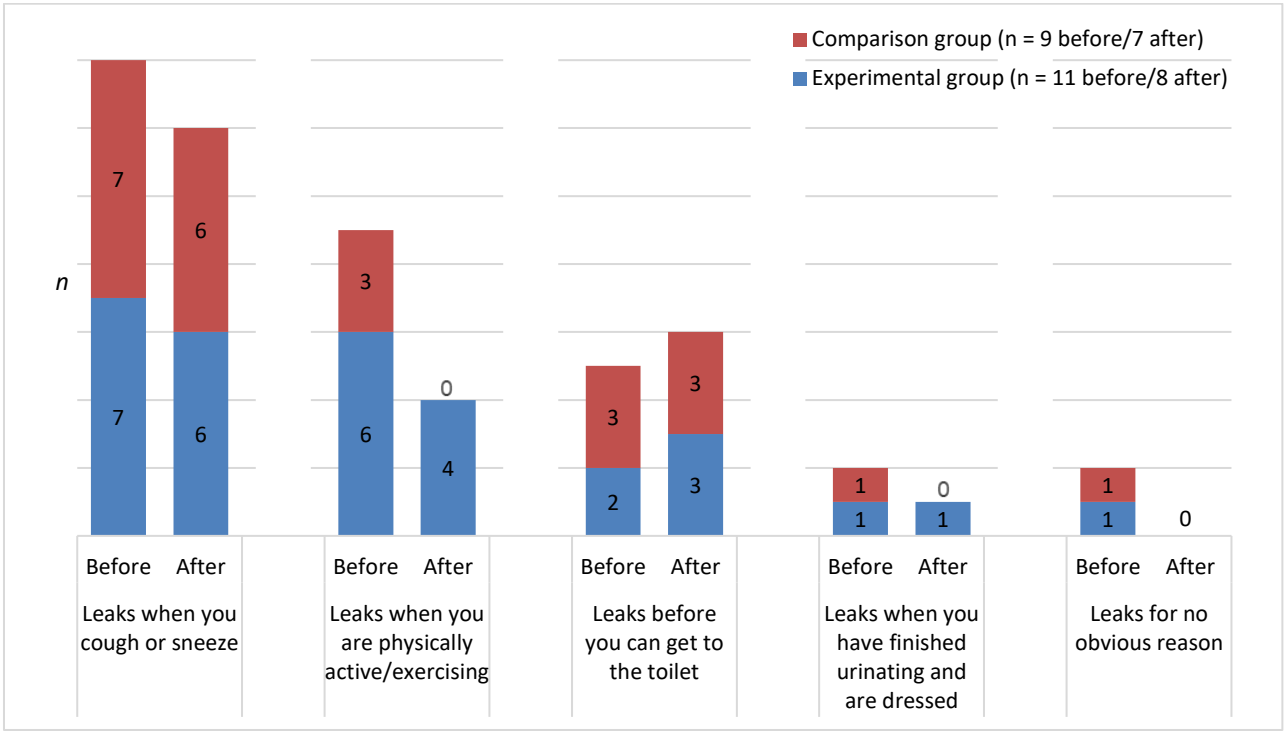
intervention group, indicating that incontinence during being physically active/exercising was the form of incontinence which decreased most in both groups.

Table 16 ICIQ-UI SF sum score[a]

			Preintervention	Postintervention
Experimental group (*n* before = 37, *n* after = 36)	Raw score *n* (%)	0 4-9 13, 14	26 (70.3) 9 (24.3) 2 (5.4)	28 (77.8) 8 (22.2) 0
	Mean (*SD*)		2.2 (3.8)	1.4 (3.2)
	Median (min-max)		0 (0-14)	0 (0-13)
	IQR		5	0
	UI rate *n* (%)		11 (29.7)	8 (22.2)
Comparison group (*n* = 18)	Raw score *n* (%)	0 1-9 13, 14	9 (50.0) 3-9: 8 (44.4) 1 (5.6)	11 (61.1) 1-6: 7 (38.9) 0
	Mean (*SD*)		3.0 (3.8)	1.5 (2.1)
	Median (min-max)		1.5 (0-14)	0 (0-6)
	IQR		5	4
	UI rate *n* (%)		9 (50)	7 (38.9)

Note. min = minimum, max = maximum; IQR = interquartile range; UI = urinary incontinence.
[a]Maximum = 21.

Figure 25 Qualitative ICIQ-UI SF results: forms of incontinence frequencies at trial entry and end for both trial groups



Note: Multiple answers at trial entry by 12 women, at trial end by 6 women. The ICIQ-UI SF categories "Leaks when you are asleep" and "Leaks all the time" were never chosen and thus are not included in the table. Later withdrawn participant excluded from experimental group at trial start. Two particpants of experimental group included at trial end with an ICIQ-UI SF sum score of > 0 who initially scored 0.

Qualitative PRO results

When asked at trial entry to describe any changes associated with their pelvic floor since birth (*n* = 28, at 6-24 weeks post partum), four participants reported no change, 10 described the pelvic floor as more loose or weak, five mentioned urinary symptoms, and two named both pelvic organ prolapse symptoms and sexual issues. Four participants described the pelvic floor as better than in pregnancy and six spoke about changes soon after birth which had already resolved; two could not describe the change.

After the intervention period, participants were asked about pelvic floor changes since the start of the intervention. Table 17 displays the answers by group and method of data collection (these data were gathered in the interview but also named in the online survey). Four participants of the experimental group reported enhanced pelvic floor awareness; one of these participants explicitly compared this to her pelvic floor muscle exercises in antenatal care with which this awareness had not developed.

Table 17 Pelvic floor changes named in open-ended question after the intervention period

	Experimental group	Comparison group
Interview	*n* = 36: – No change (*n* = 13) – Pelvic floor more firm (*n* = 10) – Pelvic floor enhanced (*n* = 3) – Decrease of urinary symptoms (*n* = 7) – Decrease of prolapse symptoms (*n* = 2) – Enhanced sexual sensations (*n* = 1) – Diminished haemorrhoids (*n* = 1) – Pelvic floor worse than before, pelvic organ prolapse symptoms ascribed to being out and carrying the child a lot – Pelvic floor weaker since not using the ball any more (*n* = 1, participant was four weeks after intervention end at final interview)	*n* = 18: – No change (*n* = 2); however, one of these two participants reported somewhere else that she was able to hold the long contractions longer towards the end of the intervention period – Pelvic floor more firm (*n* = 9) – Decrease of urinary symptoms (*n* = 6) – Enhanced sexual sensations (*n* = 1) – "Symptoms" (from her preintervention data identified as vaginal, urinary and anal) gone (*n* = 1) – Hitherto unknown feeling of downwards pressure during menstruation (*n* = 1) – Now able to release the pelvic floor in a controlled way (*n* = 1)
Online survey (anonymous)	14/18 questionnaires completed, 13 participants answered question about opinion on pelvic floor muscle strengthening by vibrating vaginal balls. With respect to effectiveness, – Two participants stated their pelvic floor had become stronger, – Two participants did not feel improved strength, and – One participant can imagine the pelvic floor could be strengthened with a heavier ball or more movement during use than that occurring during housework.	7/8 questionnaires completed, six participants answered question about opinion on pelvic floor muscle strengthening by exercises. With respect to effectiveness, – Three participants stated, and two of them emphasised this, that the exercises were effective.

Sixteen participants (11 in the experimental and five in the comparison group) could not answer the question on enhanced sexual sensations by intervention use as they had not or only resumed little respective sexual activity. In the experimental group, 16 (of 36) reported no changes. Of the four that did report a change, the commentaries were:

- “Maybe the pelvic floor is better circulated, firmer” (ID 19);
- “I feel the pelvic floor more” (ID 9);
- “I am more conscious of the pelvic floor, can utilise it more consciously, strengthen it more consciouly as female organ, and my husband has confirmed this; I always heard this but didn´t take it seriously” (ID 21); and
- “Ball use was the start for more relaxed sex, we were not as careful as before any more” (ID 5).

In the comparison group, five (of 18) participants reported no changes. Five participants ascribed enhanced sexual sensations to pelvic floor muscle training. One participant commented that she herself had not noticed a rise of muscle tone but her husband had. One participant could not separate the effect of the exercises from the effect of stopping breastfeeding, and one knew this effect from her past but had not experienced it in the trial intervention period. One participant found exercising itself sexually stimulating.

8.2.3 Vaginal manometry

Table 18 displays the descriptive statistics for each perineometry outcome per group before and after the interventions. It can be seen that for maximum strength squeeze pressure, mean pressure of the three strength squeezes, endurance and endurance (10)[92], the values rose within both groups; only the values for vaginal resting pressure shrank in both groups. The initial point estimates are slightly different between the groups for all outcomes.

Of the 53 within subject maximum strength measurements, 39 (73.6%) rose, 12 (22.6%) dropped, and two (3.8%) stayed the same. The participants with smaller initial values had larger percentage increases, with 13.0% of the rise of maximum strength being accounted for by initial value. The drop of the 12 reduced values was between 1% and 42%. Seven of the 12 reduced values were in the experimental group, five in the comparison group, and the two identical values were in the experimental group.

Table 19 displays the change scores per group and the change score differences as effect sizes, and Figure 26 a-c is a graphical representation of selected group change scores. For vaginal resting pressure, the experimental group displays a small fall descriptively but a rise of 0.1 in the model; this difference between descriptive and model data is explained by the

[92] Operationalised with a maximum of 10 seconds (see section 5.4.2).

number of included observations. Vaginal resting pressure change score difference between the groups is 2.8 cm H_2O (95% CI [-2.2, 7.9]) (indicating a higher rise for the experimental group). The experimental group also displays a higher rise for maximum strength squeeze pressure (change score difference 4.6 cm H_2O, 95% CI [-0.3, 9.4]) and the mean of the three strength squeezes (change score difference 3.1, 95% CI [-2.2, 8.4]). It displays smaller rises for endurance (change score difference for endurance -1.9 cm H_2O, 95% CI [-8.4, 4.7], and for endurance (10) -0.4 cm H_2O, 95% CI [-2.0, 1.1]).

For maximum strength squeeze pressure as potential future primary clinical outcome, a PP analysis was performed to see whether there was a possible effect according to the reported adherence to the intervention protocol. This PP analysis included 22 participants in the experimental and 9 in the comparison group; the change scores predicted from the statistical PP model are 8.2 cm H_2O (*SD* 10.7) and 2.7 cm H_2O (*SD* 5.2) respectively. With a change score difference of 5.5 cm H_2O, 95% CI [-0.4, 11.3], the PP analysis delineates a directionwise identical but larger effect size than the ITT analysis, thus supporting the mITT analysis.

Table 18 Descriptive results for perineometry

Outcome	Experimental group			Comparison group		
	n pre/post	Pre[a]	Post	*n* pre/post	Pre[a]	Post
		Mean (*SD*)			Mean (*SD*)	
Vaginal resting pressure (cm H_2O)	37/36	25.7 (10.3)	25.4 (9.0)	18/18	29.1 (9.1)	26.4 (6.9)
Maximum strength squeeze pressure (cm H_2O)	37/36	22.1 (14.8)	30.1 (20.5)	17/18	24.9 (14.3)	27.2 (15.3)
Mean pressure of three strength squeezes (cm H_2O)	37/36	19.4 (13.9)	26.5 (18.2)	17/18	21.6 (13.7)	24.6 (14.4)
Endurance (seconds)	34/35	13.3 (9.2)	18.7 (15.9)	16/18	10.3 (6.8)	15.6 (10.5)
Endurance (10) (seconds)	34/35	7.7 (3.4)	8.1 (3.4)	16/18	7.4 (3.6)	7.7 (3.8)

Note: pre = preintervention, post = postintervention.
[a]First nine preintervention values (in total, irrespective of group) collected without decimal places.

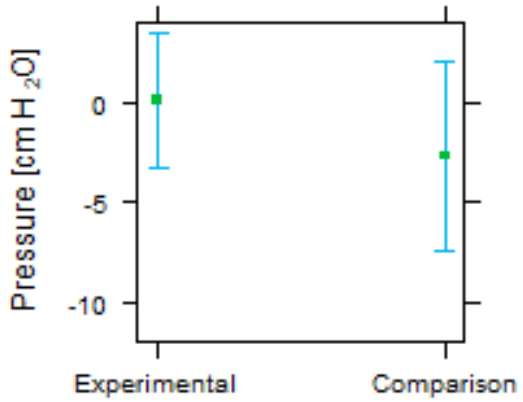
Table 19 Perineometry: modeled change scores by group and effect sizes

Outcome	Experimental group		Comparison group		Effect size
	n pre/post	Change score (*SD*)	*n* pre/post	Change score (*SD*)	Change score difference [95% CI]
Vaginal resting pressure[a] (cm H_2O)	37/36	0.1 (11.4)	18/18	-2.7 (7.0)	2.8 [-2.2, 7.9]
Maximum strength squeeze pressure[a] (cm H_2O)	37/36	7.7 (12.2)	17/18	3.1 (5.4)	4.6 [-0.3, 9.4]
Mean pressure of three strength squeezes[a] (cm H_2O)	37/36	6.8 (12.4)	17/18	3.7 (6.7)	3.1 [-2.2, 8.4]
Endurance (seconds)	34/35	5.2 (14.1)	16/18	7.1 (8.3)	-1.9 [-8.4, 4.7]
Endurance (10) (seconds)	34/35	0.6 (2.8)	16/18	1.1 (2.4)	-0.5 [-2.0, 1.1]

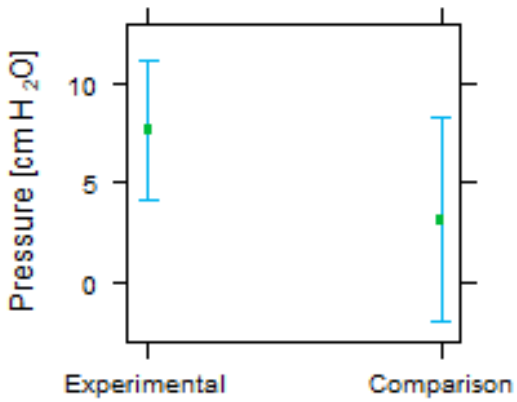
Note: pre = preintervention, post = postintervention.
[a]First nine preintervention values (irrespective of group) collected without decimal places.

Figure 26 Perineometry change scores per intervention group (means with 95% CIs)

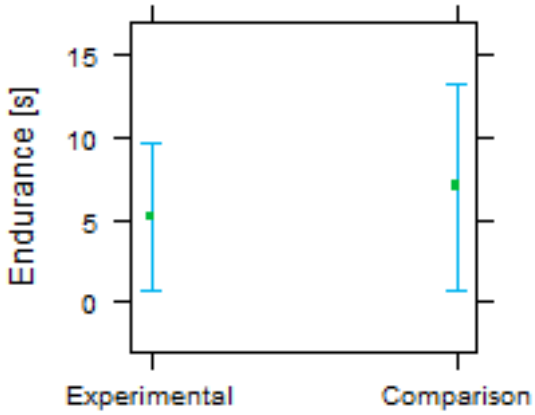
(a) Vaginal resting pressure
(*n* = 36 experimental/18 comparison)

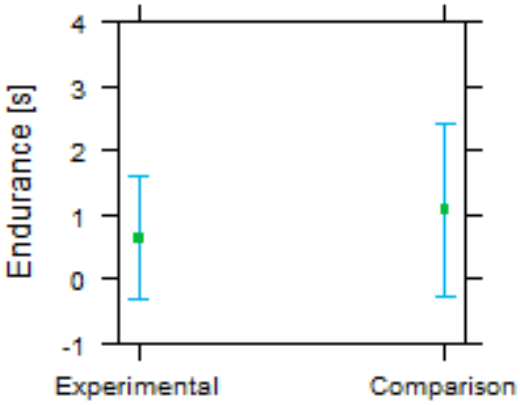


(b) Maximum strength pressure
(*n* = 36 experimental/17 comparison)



(c) Pressure endurance (in seconds)
(*n* = 34 experimental/16 comparison), measured values (left) and endurance (10) (right)





8.3 Harms

From the 55 participants included in the clinical analysis, a nonadherent participant in the comparison group who never performed pelvic floor muscle exercises is only considered for the outcome infection and not for all other harms items. The reason for including this particular item only for this participant is that vaginal infection as a common disease (Kent, 1991, Foxman et al., 2013) might more likely occur independently of pelvic floor muscle training than the other harms items which can be assumed to be more likely related to training.

Of the thus 55 and 54 participants considered for infection and other harms items respectively, 28 participants reported a total of 38 adverse events (not counting recurrence),

whereby 20 participants experienced one, six participants two, and two participants three adverse events. Table 20 gives a group-specific overview on kind and number of the main adverse events identified. Six events in five participants were treatment emergent adverse events (vulvovaginal symptoms/vaginal candidiasis/cystitis), one of them was recurrent (three times vulvovaginal symptoms in one participant). There was no serious adverse event as defined by the European Medicines Agency (1995). For the mostly minor adverse events, severity was not evaluated formally as not relevant, but severity was considered for vulvovaginal symptoms/infections. Except for these which in one case led to participant withdrawal and in all other cases to an intervention-free trial interval, the adverse events did not impact on following the intervention protocol. As in the effect results sections, the uneven 2:1 sample distribution must be considered when interpreting results for harms.

Table 20 Main group specific adverse events

Adverse event	**Experimental group** (*n* = 37) *n* % [95% CI] *CAN* (method of calculating CI)	**Comparison group** (*n* = 18) *n* % [95% CI] *CAN* (method of calculating CI)
Vulvovaginal symptoms/ candidiasis	5 13.5 [4.5, 28.8] 4.32 (exact) Recurrent: 1	0[a] 0 [0, 16.7] 0 (rule of 3)
Cystitis[b] (on top of candidiasis)	1 2.7 [0.1, 14.2] 0.97 (exact) Led to withdrawal	Likely to be 0, but screened only for “vaginal candidiasis or other vaginal infection”
Muscle soreness	5[c] 13.9 [4.7, 29.5] 4.31 (exact) Mentioned recurrence: 1	3[d] 17.6 [3.8, 43.4] 2.47 (exact) Mentioned recurrence: 2
Local discomfort/pain	12[c] 33.3 [17.9, 48.7] 8.0 (approximate) Most items imply recurrence	4[d] 23.5 [6.8, 49.9] 3.06 (exact) Most items imply recurrence

Note. CAN = criterion for approximate normality.
[a]*n* = 17 as one value missing. [b]Active reporting only. [c]*n* = 36 as value missing for the participant who withdrew. [d]*n* = 17 as nonadherent participant not considered.

Vulvovaginal symptoms/infection

The six vulvovaginal symptom cases/infections occurred in the experimental group. The participant with the candidiasis and cystitis reported candidiasis after 11 days of ball use, did not consult her gynaecologist immediately (although she assured CO on the phone she would do so) and then developed cystitis before she finally saw her doctor who confirmed the diagnosis (according to participant). She withdrew from the trial.

Another participant developed vulvovaginal symptoms after three days of ball use and had two recurrent episodes. She was after her third birth and reported having suffered from vaginal candidiases after her first two births at the final study meeting only; when asked about vaginal infections at the initial screening, her former postpartum experiences had not crossed her mind. One participant attributed her vulvovaginal symptoms to the public swimming pool as it only developed after 10 weeks of ball use and she at that time often went swimming. The remaining two reported vulvovaginal symptoms after 10 days and eight weeks of ball use respectively. The participants self-diagnosed their vulvovaginal symptoms as candidiasis and treated them themselves as such. All symptoms were cured as the trial progressed, and all women continued participation.

If the vulvovaginal symptoms were candidiases indeed, they would fit the label “uncomplicated” from a medical point of view (as defined by Sobel et al., 1998, Centers for Disease Control and Prevention, 2015). The only information that could be elicited from the participant about cystitis severity was that treatment included antibiotics. As the vulvovaginal symptoms and candidiasis/cystitis led to withdrawal of one participant and needed treatment in all, these could be considered severe adverse events in relation to using a vaginal device for the purpose of pelvic floor muscle strengthening, even if no participant did call any adverse event severe (feedback missing from withdrawn participant).

Two events in the experimental group were not categorised as infection but shall be reported. One participant mentioned a sore spot on a labium at about three weeks after intervention start; she continued after a short intervention pause. The participant with the known predisposition to herpes mentioned two episodes of incipient herpes which she attributed to stress, she paused ball use on both occasions.

Muscle soreness and local discomfort/pain

Pelvic floor or abdominal muscle soreness was mentioned by eight participants. Five were in the experimental group (all pelvic floor muscle soreness; one of them “possibly”) and three in the comparison group (one abdominal muscle soreness, one pelvic floor and abdominal muscle soreness; one “probably” at unknown site). One participant wanted to know how

pelvic floor muscle soreness feels and over-exercised on purpose, her resulting pelvic floor muscle soreness was not categorised as adverse event. One participant mentioned muscular tension but as it was not clear at the analysis stage if this was meant as success (strengthened muscle) or adverse event, it was not counted as adverse event. Local discomfort or pain other than muscle soreness was reported by 15 participants whose statements can be found in Table 21.

Table 21 Participants' statements on local discomfort or pain other than muscle soreness (multiple answers possible)

Experimental group (*n* = 12)	**Comparison group** (*n* = 3)
Ball insertion: – Retraction cord uncomfortable at ball insertion – First insertion difficult (vaginal tightness) as first time since birth that something as large was inserted – Initially pain/discomfort at ball insertion – Inserting the ball was a bit uncomfortable (at the beginning) – Ball insertion a bit uncomfortable at the beginning (scar?) – Insertion without lubricant uncomfortable – Injured once at ball insertion without lubricant **During ball use:** – Feeling retraction cord – Slight burning sensation with ball use since on contraceptive pill – Burning sensation (from vulvovaginal symptoms?) – At the beginning: not so comfortable during use, can feel ball and it feels good to take it out **Removing ball:** – Removing the ball was a bit uncomfortable (at the beginning) – Burning sensation when taking ball out in 30-40% of occasions **After ball use:** – Dragging sensation at the right side [of pelvic floor?] after too long a use	**During exercising:** – Sometimes some burning sensation at vaginal entrance during exercising – Pain in lower abdomen at ovulation, uncomfortable during period – Now and then period like pain during exercising

Bleeding

Nonphysiological vaginal bleeding, explored only in the experimental group, was described by four participants. One of them reported slight vaginal bleeding at the first times of ball use. One participant had her gynaecologist confirm that the bleeding did not stem from ball use, one had slight bleeding at three occasions and questioned herself that the bleeding was caused by the ball, and one was not sure in her recall of bleeding.

Other adverse events

Six participants mentioned other adverse events. In the experimental group and after probing, three named vaginal discharge: One attributed this possibly to the lubricant, one “maybe” had some yellowish discharge on the ball (always), and one had one episode which might have been from ovulation (this participant also had vulvovaginal symptoms). One participant mentioned that her haemorrhoids had worsened. In the comparison group, one participant found the exercises tiring at the beginning, and one described that the exercises had kept her from falling asleep when performed before going to bed.

Summary

Knowledge of participants’ clinical experience with a vibrating vaginal ball could be increased. Nine felt the ball slip out at least sometimes. Although some participants lost the ball at some point while wearing it, holding it in the vagina for the intended timespan was usually possible for all but one participant. A quarter of the participants did not perceive the inserted ball, more than half did not feel a reaction of the pelvic floor, although some felt an enduring contraction and almost a quarter performed voluntary contractions to keep the ball in. The ball’s vibrations were mostly felt with stronger movements only or not at all. Ball use enhanced pelvic floor awareness in four participants.

Participant reported pelvic floor muscle strength showed a change score difference of -5.1% (95% CI [-18.0, 7.8]) between the groups (favouring the comparison group); participants in the comparison group were also 1.2 times (RR 1.23, 95% CI [0.95, 1.60]) more likely to report increased pelvic floor muscle strength. Descriptive symptom changes including the ICIQ-UI SF supported this larger increase in the comparison group. Pelvic floor muscle performance measurements by perineometry showed a high variability and wide effect size CIs. The effect size and 95% CI for maximum strength was 4.6 [-0.3, 9.4]. In the open-ended interview questions, more positive pelvic floor changes were ascribed to the comparison than to the experimental intervention. The preliminary harms analysis revealed more infections and vulvovaginal symptoms in the experimental group. Likewise, local discomfort/pain was

named more often in this group; minor adverse events were muscle soreness (both groups) and vaginal bleeding (experimental group).

The following last results chapter finally reports participants' experiences with and opinion on the trial.

9 RESULTS IV – PARTICIPANTS' EXPERIENCES WITH AND OPINION ON THE TRIAL

This last results chapter specifically addresses the research question on participants' perspectives on and experiences with the trial, which also contributes to the overall research question on how best to prepare and perform a future full RCT.

Participants' experiences with and opinion on trial participation were asked in 28 long interviews and 26 online survey questionnaires at the end of their participation period. The first 18 participants from the experimental and first 10 from the comparison group were interviewed. The remaining participants (18 in the experimental and eight in the comparison group) were sent a link to their respective online survey questionnaire; of these, 14 and seven respectively completed the questionnaire, mirroring the 2:1 group allocation and resulting in a survey response rate of 80.8% (21/26). Feedback on trial experience and opinion on trial thus was available from 49 participants.

Asked for their experience as a study participant, 44 of 49 participants (89.8%) gave a positive statement: 21 called it interesting/exciting, eight felt they had been well cared for, six mentioned it had enhanced their motivation to do something for the pelvic floor, and three found it uncomplicated. Amongst other positive remarks were being happy to have participated, the feeling to contribute to something, to have "proudly told it around" (ID 6), and that the husband as a "team player" had reminded the participant to use the ball (ID 21). One participant stated that she would participate again and can recommend it to other mothers. Among the six negative statements (12.2% of 49 participants, one in the comparison and five in the experimental group) were time and adherence difficulties ("participation stressful") and vulvovaginal symptoms. Three participants in the experimental group expressed relief that the intervention period was over.

Forty participants (of 49, 81.6%) named advantages of trial participation. These were e.g. the motivation to do something for the pelvic floor (*n* = 21), learning/experiencing something new (*n* = 17), a (possibly) stronger pelvic floor (*n* = 8), engagement with the topic (*n* = 3), the pelvic floor measurement (*n* = 3), and no costs (*n* = 2). Six respondents (of 49, 12.2%) did not see any advantage in their study participation. Among the seven disadvantages named were "additional obligation" (*n* = 4), feeling guilty for nonadherence, newly emerged worries about the pelvic floor, and the constraint of not being able to attend a postnatal exercise class (*n* = 1 each). Nineteen participants gave suggestions to enhance the trial which are summarised in Box 8 and concern various trial aspects.

There mostly was congruence between the trial groups, but in the experimental group, the trial experience more often was called good or interesting (28/32 [87.5%] vs 9/17 [52.9%] in the comparison group). Motivation to do something for the pelvic floor (15/17 [88.2%] vs 12/32 [37.5%]) and a stronger pelvic floor (4/17 [23.5%] vs 4/32 [but 2 of these “possibly"] [12.5%]) were more often mentioned in the comparison group. While more disadvantages were named in the interviews (6/28 [21.4%] vs 1/21 [4.8%] in the online survey), more negative comments on the experience were given online (5/21 [23.8%] vs 1/28 [3.6%] in the interviews); satisfaction with care was also more often mentioned online (7/21 [33.3%] vs 1/28 [3.6%] in interviews). The fact that feedback from the online survey showed basic congruence with interview responses seems to indicate that interview participants gave honest answers and did not withhold critical comments.

Box 8 Participants’ suggestions to enhance the trial (*n* = 1 mention unless indicated otherwise)

Experimental group:

- Modifications of adherence sheet:
  - Note minutes of ball use aimed for on adherence sheet
  - Clarify sentence about training maximum on adherence sheet
  - Delete weekday names from heading line in adherence sheets (*n* = 2)
  - Do not ask for minutes of use but for ticking adherence on form
  - Provide a legend to denote “illness”
  - Find a solution to denote sitting times
- Longer than 30 minutes of ball use per day as 30 minutes is too short and cannot be integrated into a day with children (*n* = 2)
- To use a heavier ball weight according to preintervention measurement
- Provide soap samples for ball cleaning during travelling
- Provide timer to indicate end of 30 minutes
- The participant who could not hold the ball inside had wished more support on how to successfully keep it in

Comparison group:

- A more thorough personal introduction to the pelvic floor muscle exercises with a professional
- Fewer exercises (less often per day or fewer per block) (*n* = 2)

Not group specific:

- Regular automatic e-mail reminders to support adherence
- Electronic application instead of paper sheet to document adherence
- Enquire subjective perception in pelvic floor questionnaire in more detail
- Better instruction on how to perform an enduring pelvic floor muscle contraction for perineometry measurement
- Take perineometry comparison values before pregnancy

When given the opportunity for additional comments online, six respondents in the experimental group contributed. Three thanked for participation, one is excited to hear about the results, one called it an interesting experience in spite of only a “small result in terms of pelvic floor muscle strength” (online ID experimental 1), and one wished CO all the best. In the comparison group, the one respondent thanked and stated that the exercises will accompany her all her life. Additional interview comments of relevance here were an interest in how the trial went for the other participants and in own and trial results.

Summary

Participants’ experiences with and opinion on trial participation were mainly positive and encouraging, as e.g. 21 participants found participation interesting/exciting and six named their enhanced motivation to do something for the pelvic floor. Among the six negative statements (one in the comparison and five in the experimental group) were time and adherence difficulties and genital symptoms. Asked for disadvantages, additional obligation, feeling guilty for a lack of adherence, or newly developed worries about the pelvic floor were mentioned among the seven namings. Participants gave suggestions to enhance the trial, e.g. regular reminders to support adherence, electronic adherence documentation, or soap samples for ball cleaning during travelling. The next chapter draws together the results presented in the last four chapters and discusses them.

10 DISCUSSION OF RESULTS CHAPTERS I-IV

The feasibility results of the past four results chapters comprise process, management, resources and clinical results, and participants' experiences with and opinion on the trial. This discussion chapter first presents a summary of the feasibility trial results for each of these feasibility areas. Subsequently, these specific results are discussed.

10.1 Summary of feasibility trial results

Feasibility trial results for each feasibility area are summarised in the following four sections.

10.1.1 Process results

Table 22 presents the feasibility criteria results. This shows that Criteria 1 (recruitment), 2 (timely initial measurement) and 4a (keeping to assigned group) were fulfilled. Criteria 3, 4b and 5 were not fulfilled. However, the aimed for rate of at least 90% for Criterion 3 (timely intervention start) lies within the 95% result CI upper limit of 92.0% and was almost reached after 7 instead of the planned 4 days. Likewise, although the final data collection (Criterion 5, targeted minimum rate 80%) was completed within 2 weeks after intervention end by only 66.7% of the participants, its upper 95% CI limit of 79.3% almost reached the targeted rate; after 3 weeks, the aimed for rate was nearly reached and well within the result 95% CI. Both Criteria 3 and 5 results thus are acceptable as tolerable from a clinical point of view but point to the necessity of strategies to enhance the respective processes in a full RCT.

Even with the use of wider than planned adherence criteria, the adherence rate (Criterion 4b) reached an overall maximum of 60.4% only, with even its upper 95% CI limit of 73.6% lying clearly beyond the targeted rate of at least 80%. Although the upper 95% CI limits in the intervention groups reached 78.6% and 78.9% respectively, these high values only stem from their large width (due to small group size) of surrounding a rate of 62.9% in the experimental and of 55.6% in the comparison group. This shows that adherence was the least feasible trial process and that there is a need for strategies to enhance adherence in a full RCT.

Table 22 Summary of feasibility criteria results

Criterion target and result	Fulfilled
(1) Target: ≥ 10% of eligible persons give consent to participate in the trial. Result: 48.3%, 95% CI [39.2, 57.4] (researcher level)	Yes
(2) Target: ≥ 90% (95% CI [80, 100]) of participants attend the first pelvic floor muscle measurement within 3 weeks of consenting to take part in the trial. Result: 92.9%, 95% CI [86.2, 99.6]	Yes
(3) Target: ≥ 90% (95% CI [80, 100]) of participants start with the intervention within 4 days of the initial pelvic floor muscle measurement by perineometry. Result: 81.8%, 95% CI [71.6, 92.0] Within 7 days: 89.1%, 95% CI [80.9, 97.3]	No
(4) Target: ≥ 80% (95% CI [70, 90]) of enrolled participants a) keep to the assigned intervention group and b) adhere adequately to the intervention. Result: a) 100%, 95% CI [93.4, 100] b) 47.2%, 95% CI [33.8, 60.6] minimum or 60.4%, 95% CI [47.2, 73.6] maximum	Yes No
(5) Target: ≥ 80% (95% CI [70-90]) of enrolled participants have the final data collection within 2 weeks of ending the intervention. Result: 66.7%, 95% CI [54.1, 79.3] Within 3 weeks: 79.6%, 95% CI [68.9, 90.4]	No

Whereas the methodology outlined in the research protocol (Appendix N) was the plan of this feasibility trial, the content of the methods chapter delineated the plan's implementation. For different reasons which were clarified in the methods chapter or discussed in the respective results sections, a number of modifications to the research protocol were applied during the course of the trial. An overview of these changes, with a differentiation between those that were required only for the feasibility trial and those which should be kept for a main RCT, is provided in Table 23.

Table 23 Modifications of trial processes from feasibility trial protocol to feasibility trial execution

Area	Item	Change (FEAS): denoting change only for feasibility RCT because of feasibility context (FULL): denoting change for feasibility and future full RCT
Recruitment	Selection criteria	Women exceptionally also accepted without six weeks postpartum check (FULL) Deleted: Criterion *Termination of participation will be recommended to a participant when there is vaginal infection* (FULL)
	Recruitment routes	Deleted: AKH Vienna postnatal ward and website of Vienna Midwifery Centre (FEAS) Added: Magazine *Eltern* and parenting centre of the City of Vienna (FEAS)
Data collection	Forms	Modifications of all except three process and data collection forms, summarised in detail in Appendix FF (FULL)
	Perineometry appointment	Introduced: Text message to be sent the day before the measurement appointment (FULL)
	Adverse events	More than the three planned calls made until participant could be reached (FULL) No adverse event collection in final online survey but question about adverse events at final call (FULL)
	Participants' perspectives and experiences	Introduced: Initial and final short structured interview for second half of participants (FEAS)
	Online survey questionnaire	Reminder at final call and per e-mail instead of planned text messages, two and four weeks later respectively (FULL)
Data processing and analysis	Data processing	Precoding of forms and SPSS codes optimised (FULL) Programme *VeraCrypt* instead of *truecrypt* (FULL) More than one data validation method performed (FULL)
	Feasibility analysis	Additional strategies applied to analyse feasibility (FEAS)
	Adherence calculation	Newly developed wider adherence criteria used (FULL)
	Effect size calculation	No sensitivity analysis for potential confounding variables and missing data (FEAS)
	Harms analysis	Application of rule of three had not been planned (FULL)

The trial process findings further show the following details:

- Active and passive recruitment routes acquired participants, whereby the active routes were more successful. Recruitment success was higher when personal contact with the recruitment professionals was established.
- Women's interest in trial participation was high as shown by an overall recruitment rate of at least 48.3% of eligible women at the researcher level.
- The most frequently named barriers to participation were the aversion to use a vaginal ball, lack of time, travel distance to measurement site, and that women were exercising at home.
- The applied selection criteria only underwent minor amendments. The sample was not representative for the target population.
- While there was a high preference for the experimental group, randomisation worked smoothly. The completion rate was high, with only one participant exclusion for ineligibility and one withdrawal.
- The data collection processes generally ran smoothly and the collection tools worked effectively; both only need(ed) minor amendments. The most challenging task was to organise the perineometry appointments.
- Suitable PROMs were developed (further) and tested.
- Piloting of trial forms contributed to their evaluation and development.
- The interventions were feasible and acceptable to participants. Reasons for the overall rather low adherence rate were in both groups having to think about doing the intervention, and in the experimental group that 30 minutes were difficult to adhere to. Ball handling was reported as easy.
- PPI positively influenced the trial, particularly by enhancing CO's confidence that the project would be feasible and of interest to women, and by checking trial documents for participants. The few pelvic floor muscle measurements by a male assessor and the paper version of the online survey questionnaire (both PPI participants' considerations) were not topics in this feasibility trial but might be of relevance in a full trial.

10.1.2 Management and resources results

Human, organisational and data management, with few minor exceptions, worked smoothly in this project. Suggestions regarding organisational trial management are a perineometry venue with an electronic device to access the online data collection form (to guarantee data protection), and a locker to store the Peritron and transport tickets; further, for data management, elaborated data file and document naming/numbering and variable coding.

Staff, time and resources needed for this feasibility trial could be calculated, which can serve as a basis for budget calculation for a full RCT.

10.1.3 Clinical results

Data on participants' clinical experience with a vibrating vaginal ball could be collected and the theoretical basis of the experimental intervention could be increased. This revealed that a quarter of the participants did not sense the presence of the inserted ball, more than half did not feel a reaction of the pelvic floor, nine felt the ball slip out at least sometimes, and almost a quarter of the participants performed voluntary contractions to keep the ball in. The ball's vibrations were often felt with stronger movements only or not at all.

All pelvic floor outcomes changed in the same direction in both intervention groups. Their rise was expected since self-improvement of pelvic floor muscle strength over time is expected to occur in postpartum women irrespective of an intervention; the only fall (of vaginal resting pressure) was unexpected and is unexplainable. The effect result for participant reported pelvic floor muscle strength showed a higher rise (with wide CI) in the comparison group. This is supported by the result for postintervention Question 2 where participants rated their pelvic floor more often as "enhanced" in the comparison group. Accordingly, the PRO symptom values shrank more in the comparison group for all symptoms (except for stool incontinence which was not comparable). Furthermore, the ICIQ-UI SF sum score mean reduction and the decrease of the urinary incontinence rate was larger in the comparison than in the experimental group although initial rates were higher in this group. When asked about pelvic floor changes in the interview, beneficial changes were also reported by more participants in the comparison than in the experimental group, as was the case with a positive influence on sexuality. The (mostly descriptive) PRO results in this trial therefore favour pelvic floor muscle training.

The calculated perineometry effect sizes are small and have wide CIs. As vaginal resting and strength pressure favour the experimental intervention, they support the PRO findings only partially. On the whole, the trial's preliminary clinical results show that, when comparing hold and vibration use of a vibrating vaginal ball of 28 g with pelvic floor muscle training at home, there was a tendency to favour the comparison group; this effect however is uncertain, and the wide CIs also indicate the potential for an effect in the opposite direction.

Harms screening revealed more vulvovaginal symptoms in the experimental group, namely 13.5% (95% CI [4.5, 28.8]) vs none (upper 95% CI limit 16.7%) in the comparison group. The only case of cystitis happened in the experimental group. With 33.3% (95% CI [17.9, 48.7]), participants in the experimental group also named local discomfort more often than participants in the comparison group with 23.5% (95% CI [6.8, 49.9]).

10.1.4 Participants' experiences with and opinion on the trial

Participants' experiences with and opinion on the trial were mainly positive and encouraging; many called participation interesting or exciting. Time and adherence issues and vulvovaginal symptoms were among the few disadvantages named. Participants provided a number of suggestions to enhance the trial.

10.2 Discussion of specific feasibility results

In the following, the results about the specific trial feasibility areas are discussed. As with results reporting, this starts with the process issues, goes over trial management, resources and the clinical aspects, and ends with the participants' experiences with and opinion on the trial.

10.2.1 Trial processes

The process of recruitment is discussed first, followed by the sample description and participant flow. Further process results cover the trial forms, the data collection, and the trial interventions.

Recruitment

Lack of access to recruitment within the NHS may have had to do with CO's lack of familiarity with the British health care system and lack of personal professional contacts in London. It was also realised during the project that vibrating vaginal balls are not well known in the UK and the topic therefore may not have been intriguing enough to UK colleagues. At the postnatal wards of the AKH Vienna, the high workload was the reason named why trial recruitment would not be supported by nursing staff, which however might be combined with a lack of awareness and budget for midwifery research. Not being allowed to place leaflets at the parenting centres of the City of Vienna was an administrational issue; as the reason was that private persons are not allowed to put out leaflets, it could, for a full trial, be argued with the institution to which the researchers will be affilitated.

The reasons that no contact dates for women were obtained via medical doctors at the postnatal wards of the AKH Vienna might be manifold. There is a high fluctuation of resident doctors discharging women and thus, information about the study may have got lost; there was no incentive for supportive doctors; the hospital doctors did not have a personal professional relationship with CO like the personally approached obstetricians and midwives; and medical doctors might lack awareness for midwifery research. In the delivery suite, the number of women discharged home (and not to the postnatal ward) is small (2016: 1.5 % (Allgemeines Krankenhaus der Stadt Wien, 2017)), and a high workload may have contributed to midwives not paying attention to trial recruitment.

Half of the out-of-hospital obstetricians approached did not respond to the invitation e-mail which might be due to the form of communication. The motivation of those who supported recruitment seemed to be an interest in the topic and appreciation of the research idea; further, they were approached by CO personally and not unknown members of a hospital team. The barriers given by the few obstetricians for not supporting recruitment lay within the circumstances of their professional practice.

The community midwives were approached by phone calls. They seemed to have been motivated by supporting a colleague in doing midwifery research, and a personal professional relationship between some of them and CO may have contributed to this. Barriers to support recruitment on the midwives' part mostly lay within the circumstances of their professional practice. Only for one midwife, recruitment support would have required too much of an effort; in a full trial, recruitment support would require less effort as feasibility data (number of women not communicating their phone number and reasons for nonparticipation) would not be collected.

At the researcher level, noneligibility accounted for about a quarter of nonrecruitment. This comprised mostly medical characteristics and (planned) attendance of postnatal exercise classes. Women who could not be reached may have reflected the habit of people not answering calls from unknown phone numbers or women with babies not having time to answer phone calls. The two women who could not be reached for a second call might not have answered the calls because they did not want to participate and felt unable to decline participation; this may also have been the case for the women who did not call back although agreed upon.

The high overall recruitment rate of at least 48.3% of eligible women and phone calls by women who were past six months post partum or attending postnatal exercise classes seems to indicate that women after childbirth in Vienna are interested in such a trial. There were even two participants who insisted on participation despite reluctance towards enrolment expressed by CO: one with recurrent genital herpes infection and one with a need for therapy because of urinary incontinence symptoms. A high interest of women was also the case in other peripartal pelvic floor studies (Nielsen et al., 1988, Chiarelli et al., 2003, Guerrero et al., 2007). The recruitment rate however was not the recruitment rate at the level of all potential participants but calculated at the researcher level with part of the women already "filtered" at the individual professional level. A recruitment rate for the individual professional level is not available but seems to have been lower (concluding from professionals' feedback). It was with this rate in mind that Feasibility criterion 1 was set at 10% of eligible women; as it was almost 50% at the researcher level, it may well have been above 10% at the professional level, thus fulfiling Criterion 1. However, Cooper et al. (2018)

cautioned to use results from external pilot (and thus feasibiltiy) trials to estimate randomisation (meaning recruitment) rates for full trials as their comparison of such rates between external pilot and their associated full trials demonstrated high variability. Also, recruitment by a larger institutional research team instead of a “PhD midwife” might differ.

When enquired, the participants’ interest stemmed to a large degree from having symptoms and seeing trial participation as a therapeutic possibility, or from being worried about pelvic floor problems in the future. This is in agreement with Brubaker et al. (2013), in whose study participants’ reasons for participation in pelvic floor studies comprised a hoped for improvement in their condition. Of further participation arguments identified by Brubaker et al. (2013), the wish to learn more about the condition can be confirmed by this trial; as in Brubaker et al. (2013), incentives and compensation were minor topics (only one potential participant in this trial asked about compensation). The high recruitment rate might also reflect Mason et al.’s (2001b) and Chiarelli et al.’s (2003) conclusion that women prefer pelvic floor muscle training with a professional over leaflets, as trial participantion offered contact with a midwife (CO) and a certain degree of supervision. Altruistic reasons for study participation were also identified by Brubaker et al. (2013) and Newington and Metcalfe (2014). The high rate of participants recruited via their midwife/obstetrician and the fact that three of them gave the recommendation by their obstetrician/midwife as a reason for participation might confirm Brubaker et al. (2013) who showed how much women trust their caregiver when deciding to join a study.

More than a quarter of the approached women declined participation. As reasons for nonparticipation, a quarter of them said they did not want to use a vaginal ball, confirming Prashar et al. (2000) who showed that vaginal devices are not acceptable to all women. Although the ball is larger than a cone, only one potential participant with a problematic scar from her birth injury declined participation upon the introduction of the ball at the recruitment visit (there was a second potential participant who declined participation when the ball was presented but the reason in her case is unclear). Donovan et al. (2016) gave treatment preference as a barrier to recruitment, as was the case here when a number of women said to prefer pelvic floor muscle training. On the other hand, the participants in this feasibility trial expressed a strong preference for the experimental group. In contrast to a few women who stated to exercise at home, others expressed concerns about their (non)adherence in the trial. Some of the reasons given for nonparticipation ressemble those found by Chiarelli et al. (2003): being too busy or back at paid work, not wanting a vaginal exam, or living too far away. Although there was a number of women for whom it was too much effort to travel to the perineometry measurement, two participants accepted a journey of 34 and 41 km (21/25 miles) to travel to the measurement site. These issues are confirmed from a researcher point

of view by Newington and Metcalfe (2014) who identified logistical issues as reasons for declining participation and suggested reducing participant burden to improve recruitment.

Women's reasons for nonparticipation were taken at face value and not explored further, but there might be different reasons behind the ones stated. For example, a few women gave the impression of not being able to decline participation directly and of giving a socially desired answer by naming reasons such as "not wanting to occupy a study place" or considering herself "not to be of use for the trial".

Eligibility criteria

Most of the selection criteria proved feasible. The inclusion criterion *Six weeks check performed* was influenced by the health care system as this is the routine close out exam for women after birth in Austria and other countries (MacArthur et al., 2003, Geist, 2005, Brodribb et al., 2013, Bundesministerium für Gesundheit, 2017, FOKUS KIND Medien, 2017) and recommended by the World Health Organization (2014). Participants' reports on their exam provided a certain reassurance about a good postpartum health status to CO, who did not physically examine the participants. Even if an exception was made for one participant (declined exam) and exceptions can be justified in a full trial, it might be sensible to keep the exam as an inclusion criterion for this reason, the more as it might also have forensic implications.

It was explained to each participant that she should try to find out whether she was able to hold the ball inside the vagina before going to the initial measurement. It was not checked whether participants had followed this instruction, but with the participant concerned it could be spotted that the ability to keep the ball inside obviously cannot be determined on a single occasion but becomes clear only over time. Her case indicates that it might be reasonable to delete the respective exclusion criterion in a full trial as it is not possible to determine at trial start whether it applies. To put the same standard on both groups (to avoid performance bias), it is recommended to delete the exclusion criterion *Termination of participation will be recommended to a participant when there is vaginal infection*.

In this feasibility trial, women with urinary incontinence symptoms were not excluded for feasibility reasons, but this needs to be reconsidered for a full RCT. Dumoulin et al. (2017) and Senat et al. (2016) recommend to treat urinary incontinence persisting at three months post partum. This high-level evidence recommendation could guide the decision to include only urinary continent women. If followed, the impact of this new exclusion criterion on the recruitment rate must be taken into account. Considering the participants in this feasibility trial shows that 10 of 21 women with urinary incontinence symptoms were under 12 weeks post partum whereas 11 were at least 12 weeks post partum. If these 21 had not been

eligible for this trial, this would have raised the noneligibility rate from 24.3% to 52.7% (39/74 checked for).

Sampling and sample

Sampling and sample issues are sample representativeness and participants' experience with and opinion on the interventions at trial start. Each of these is considered in the following.

Representativeness of the sample

Opportunity sampling is a form of nonprobability sampling (Lund Research Ltd, 2012). This means that every potential participant (here: all postpartum women in Vienna fulfilling the selection criteria) has an unknown chance of selection, and there is a lack of knowledge about whether participants differ in some way from nonparticipants (Burns, 2000). Of all sampling methods, opportunity sampling is the weakest strategy with the highest risk of selection bias and least generalisability (LoBiondo-Wood and Haber, 2013a, Indrayan, no date). Participants might not represent the population of interest as the decision to participate in the study might reflect some inherent bias in the characteristics of the participants, and those who feel strongly about the issue in question may favour a certain outcome (Moore and Notz, 2006)[93]. The results might thus exaggerate the findings (Lund Research Ltd, 2012, LoBiondo-Wood and Haber, 2013a) and not be generalisable.

It was attempted to determine sample representativeness by comparing the sample's characteristics to those of the Austrian birthing population. In clinical terms, this comparison is not adequate as the present trial only included women with vaginal delivery at term without large perineal (anal sphincter) tear who represent around 60% of Austria's birthing population[94]. While these criteria are appropriate for the purpose of this trial, this selection does in clinical terms not represent the future trial's target population[95] of all postpartum women who might want to use a vibrating vaginal ball. Also, no participant with a multiple birth was recruited. Trial participants were recruited in (and around) a single city and were limited to women being able to communicate in German. Further differences caused by the selection criteria and not analysed here might lie between the sample and the Austrian childbearing population.

[93] This is also a problem inherent in research in general as it is not possible to study people who do not consent to being studied (Meinert, 2012). Participants therefore never are a random sample (Kiene, 2001), and results from participants and nonparticipants might differ (Sackett, 1979).
[94] Caesarean section rate in 2015 was 30.2%, preterm birth rate until 36+6 weeks of pregnancy 8.0%, 3rd/4th degree tear rate 2.0% (Institut für klinische Epidemiologie der Tirol Kliniken GmbH, 2016).
[95] The population to which results are planned to be generalisable (LoBiondo-Wood and Haber, 2013a).

An accurate determination of sample representativeness for the Austrian childbearing population in terms of ethnicity[96] is not possible. According to the national statistical office (Statistik Austria, 2015), the British ethnicity framework (as in National Statistics, 2003) does not exist in Austria. A roughly similar index in the Austrian statistical registry is place of birth and citizenship of the newborn's parents (Bundeskanzleramt Österreich, 2013b, a). The freely available result of this national analysis however only grossly differentiates between EU and non-EU countries, and continents (Statistik Austria, 2018a). The published statistics on the newborns' nationality make a finer differentiation. Compared to these data, the sample is at least lacking women from Turkey (as 1.7% of newborns in Austria are of Turkish nationality) and former Yugoslavia (one participant [1.8%] from Serbia vs 3.8% of newborns of Ex-Yugoslavian [without Croatia and Slovenia] nationality) (Statistik Austria, 2017a); as further newborns (and mothers) could be of Austrian nationality but other cultural or ethnic origin, this would increase this deficit even further. The sample therefore can neither be considered representative of Austria's contemporary cultural diversity or of the target population of this feasibility trial, nor of the ethnic diversity[97] of a full RCT's target population.

Another demographic difference between the present sample and the Austrian birthing population lies in the participants' social background. Their completed education is higher than that of all Austrian women giving birth to a live[98] newborn (e.g. university degree 75.0% vs 15.2% [in 2013[99], 13.4% unknown] (Statistik Austria, 2018a)), and more women are employed (85.7% vs 63.3%, 14.3% unknown, in 2013 (Statistik Austria, 2018a)). This sample composition corresponds to the knowledge that study participants have a higher educational status (Hilde et al., 2012) and are socially better situated than nonparticipants, and that study samples lack minority ethnic groups (Martikainen et al., 2007, Weinberg and Chronic Disease Research Center Team, 2008).

The sample composition might also mirror the recruitment path via community midwives. Postpartum midwifery care in Austria is not mainstream, with only 12% of women in 2012 having made use of postpartum midwifery care at home (Österreichisches Hebammengremium, 2018c). Midwives' clients are probably socially better situated; however, the only demographical indicators that could be found to support the latter statement were collected about the women who made use of the free midwifery consultation

[96] The medical birth registry asks for country of birth of the mother's parents, whether the mother has been living mainly in Austria since her birth, and whether German is one of her first languages (Institut für klinische Epidemiologie der Tirol Kliniken GmbH, 2016). However, the data collected are not valid because these indicators have only been introduced recently and their collection is not well established yet (Institut für klinische Epidemiologie der Tirol Kliniken GmbH, 2017b). Therefore, comparison of trial data is to national statistical office birth registry data.
[97] Of interest as the pelvic floor shows variations by ethnicity (Milsom et al., 2017).
[98] The values for all births are not freely available but the rate of death births in 2016 was only 0.33% (Statistik Austria, 2017b).
[99] Most recent available year with a full dataset.

in pregnancy. They show that 35.5% of these women have a university degree and that 96.8% live in a partnership (Österreichisches Hebammengremium, 2018d).

Compared to the Austrian childbearing population, the participants also lived a healthier lifestyle. The Austrian breastfeeding rate was 67.0% at three months post partum and 51.2% at six months post partum in a survey with potential response bias (Esberger, 2007), whereas the initial breastfeeding rate in the present sample was 96.4%. The cigarette smoking rate in the sample was 3.6% compared to 14.2% in the Austrian childbearing population (data available only for 47.1% of women) (Institut für klinische Epidemiologie der Tirol Kliniken GmbH, 2017a). Further clinical differences are shown in Table 24.

In order for studies to be of use for as wide a population as possible, the participants should be representative so that results can be generalised. In this trial, women's characteristics that need to be represented more often are younger age, lower socioeconomic status and different ethnicities. Therefore, techniques to make the sample more representative of birthing women need to be applied in a full trial. An example is to provide language support for individuals who do not speak the study language (well enough) (Newington and Metcalfe, 2014).

Table 24 Clinical differences between sample (*N* = 56) and women giving birth in Austrian hospitals[100] (Institut für klinische Epidemiologie der Tirol Kliniken GmbH, 2017a, 2018)

Characteristic		Sample value		Austrian comparison value
Age (years) Mean (*SD*)		33.3[a] (4.4)	>	30.6 (5.3) (2015)
Parity *n* (%)	P I P II P ≥ III	25 (44.6) 28 (50.0) 3 (5.4)	< > <	48.4% 35.0% 18.8%
BMI[b] Mean (*SD*)		22.4 (2.6)	<	23.8 (4.8) (in 2015, 14.9% missing)
Vaginal birth mode *n* (%)	Spontaneous Ventouse	51 (91.1) 5 (8.9)	> ≈	62.1% 7.2%
Birth injury *n* (%)	Episiotomy Tear 2nd degree	4 (7.1) 5 (8.9)	< <	14.7% of vaginal births 14.2% (2015) of vaginal births
Birth weight newborn (g) Mean (*SD*)		3399 (398)	≈	3400 (450) (2015, all newborns ≥ 37+0)

[a]A slightly higher age is expected in the sample as age rises between birth and the timepoint at which participants entered the trial.
[b]The sample BMI is based on post partum weight whereas the Austrian BMI is calculated from weight at the beginning of pregnancy.

Participants' experiences with and opinion on the interventions at trial start

Almost all participants claimed experience with pelvic floor muscle exercises at trial entry (similar to e.g. in Hilde et al. (2012)). However, when asked for the context, they named, along with e.g. a postnatal exercise class by a midwife, diverse exercising techniques which per se are not or do not contain evidence-based pelvic floor muscle training (Pilates, Yoga, Cantienica®, Kanga, and use of a vaginal ball). It therefore cannot be assured whether all participants had experience with pelvic floor muscle exercises in the sense of repeated pelvic floor contractions. However, after the reference birth, most participants had received information on pelvic floor muscle exercises (including repeated contractions) and about three quarters (of all participants) had exercised, with almost half of them at least two to three times a week at trial start (this rate might be slightly incorrect as the pelvic floor muscle exercises categories were created after data collection when the originally planned categories did not prove optimal). This knowledge and recommendation of pelvic floor

[100] The Austrian medical birth registry only comprises women giving birth in Austrian hospitals. The approximately 1% of homebirths are not included (Institut für klinische Epidemiologie der Tirol Kliniken GmbH, 2017b).

muscle exercises corresponds to the findings of other studies, in which e.g. 64% (Fine et al., 2007) and 86% (Mason et al., 2001b) of women post partum had received instruction. Participants mainly considered pelvic floor muscle exercises important and useful. However, they raised motivational difficulties, and four participants called the exercises “boring” and “cumbersome”.

Participants had exercised with or without supervision/classes and with instructors of different professions (midwives, physiotherapists, fitness trainers) as was the case in other studies (Mason et al., 2001b, Fine et al., 2007, Hilde et al., 2012). Although the majority of participants reported to have performed pelvic floor muscle exercises in the past, they may not have performed correct contractions or received feedback on their contractions as reporting to exercise is not the same as training effectively (Hilde et al., 2012). Correspondingly, the issue of not being sure about correct exercise performance was raised by two participants.

About two thirds of the participants had heard of vibrating vaginal balls and one of eight had experience with this device. Their opinion on vibrating vaginal balls was to a great extent positive, with a fair number of neutral and a few sceptical women.

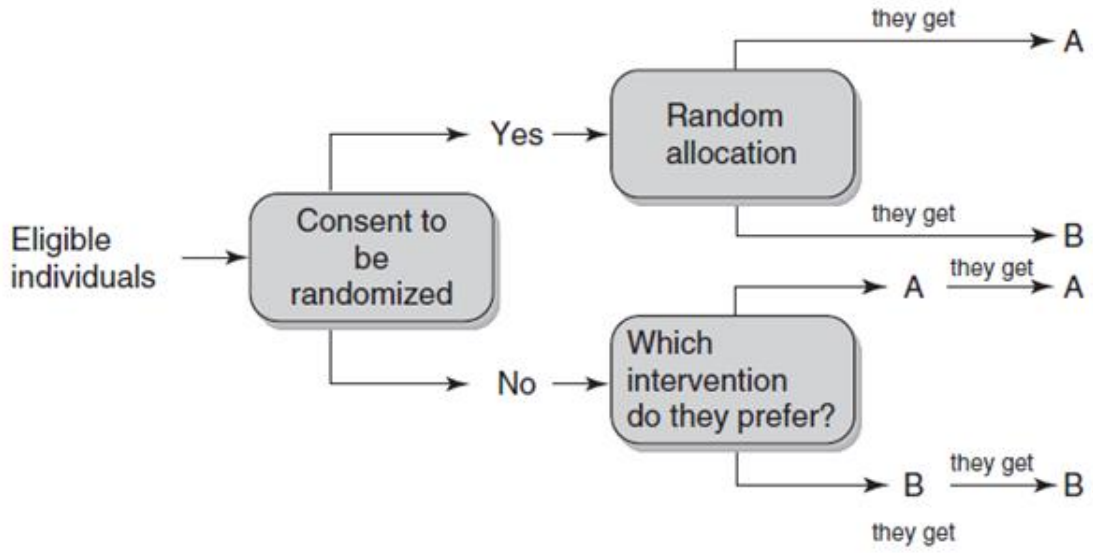
Randomisation and completion

For practical purposes, this feasibility trial was planned such that the application of the exclusion criteria *Retention of ball is impossible* and *Inability to perform the proposed procedures* was determined after randomisation. This led to one case for each criterion in which the participant was (properly) randomised before her ineligibility for inclusion could be confirmed. During the analysis stage, it became clear that this poses a difficulty in that excluding already included participants from the effect analysis does complicate an ITT analysis. Therefore, Fergusson et al. (2002) recommended that randomisation should be postponed until eligibility can be confirmed.

To overcome the problem of detecting noneligibility and excluding women after inclusion and treatment assignment, a physical examination by the researcher at the information/consent/initial study meeting could provide information on women’s ability to squeeze their pelvic floor muscles and confirm this eligibility criterion before including them in a full trial. If necessary, the examination could at the same time help women to locate and learn to contract their pelvic floor muscles (although women do not get this kind of feedback with usual care). This suggestion however would mean a third vaginal examination for the participants; to avoid this, recruitment could be tied to the six weeks post partum check and the ability to contract determined at the respective routine vaginal exam (if such an exam is performed in the trial setting).

Whereas a frequent reason for declining participation was not wanting to use a vaginal ball, participants showed a high preference for the experimental group. This became clear when asked directly, and also from participants' comments after randomisation. When potential participants have a high preference for a group, it is worth considering a preference trial design to elicit the effect of preference. Of different preference design versions presented by Jadad and Enkin (2007), the comprehensive cohort design reproduced in Figure 27 seems to be most suitable for a future RCT of the present feasibility trial. It must be borne in mind, however, that the provision of a preference study might also discourage RCT participation (Beard et al., 2015, Donovan et al., 2016).

Figure 27 Comprehensive cohort preference trial design



Source: Jadad and Enkin (2007, p. 23)

After randomisation, the participants indicated a high motivation to perform their intervention during the trial. This corresponds to the high recruitment rate and might stem from the most frequent reason given as motivation for study participation, namely to do something for the pelvic floor. Present pelvic floor impairment or hoping for an effect might also be motivators, corresponding to the continence status of 21 participants whose score indicated at least some incontinence. This result is similar to Mason et al. (2001a), where women were motivated to exercise if they had a current problem with stress incontinence or to prevent future symptoms from occurring. In their study, knowing an incontinence sufferer was another motivational factor which in this study was mentioned by one participant. According to Cammu et al. (2004), motivation also depends on the personality and enthusiasm of the physical therapist, whose role in the case of this feasibility trial was fulfilled by CO. A

limitation to the motivation results is that the validity and reliability of the motivation evaluation methods have not been formally tested.

The completion rate in this trial was high, and the attrition rate accordingly lay at the lower end of the approximately 6-19% in preventive postpartum pelvic floor muscle exercise studies (Sleep and Grant, 1987, Mørkved and Bø, 1997, Chiarelli and Cockburn, 2002, Ewings et al., 2005, Hilde et al., 2013a) or the group-specific 22-63% in Wilson and Herbison's (1998) cone trial. This may have been co-caused by the midwife-client like relationship between CO and the participants as a good rapport between the study coordinator and the participant can help minimise the dropout rate and contribute to a higher completion rate (Thoma et al., 2010, Clement, 2017); similarly, participants gave positive feedback about the good quality of care received from the perineometry assessors. Reasons for attrition were the exclusion for detecting noneligibility after inclusion (discussed above) and one withdrawal[101] for adverse events. In accordance with the fact that perinatal research in general deals with a highly mobile population (Horak et al., 2014), a few participants moved during the trial. As Austria is a relatively small country, participants were still able to travel to the final measurement and did not leave the trial, but this might not be possible in larger countries (in Wilson and Herbison (1998), 3.2% of the participants dropped out of the trial because they had moved from the area). As for randomisation rates, Cooper et al. (2018) suggested that attrition rate estimations from external pilot (and feasibility) trials should in general be used with caution because of the high variability found when comparing rates between external pilot and their associated full trials.

Trial forms and data collection processes/tools

According to a feasibility trial's purpose, all trial forms for a future RCT could be tested and optimised, by this enhancing their reliability in a full RCT. On the whole, the forms were found acceptable; most of them underwent (slight) changes, and two forms were newly created. After the changes, not all the then needed information could be found on the earlier versions which resulted in a few missing data. With the exception of the restricted online access to the perineometry form, the forms' administration proved feasible. As suggested by participants, electronic forms for participants (and consequently for the trial team) might be created for a full trial. Appendix FF sets out in detail for which forms this might be feasible.

In the following, the results about organisation of data collection and detailed future data collection processes are discussed.

[101] In spite of the recommendation to follow up withdrawn participants to collect outcome data (Gupta, 2011), this was not possible for the one participant concerned.

Organising appointments with (potential) participants

Based on the experience gained with organising study visits, a recommendation for a full trial is to keep offering home visits, preferably before noon. A low rate of visits with (potential) participant nonattendancies or without successful recruitment must be calculated. Perineometry appointments should be offered more often than once or twice a week only as for a number of participants it was difficult to find a convenient time. Reminder text messages for appointments should be sent. A challenge at the perineometry appointments was the accompanying baby/children needing attention. Babysitting service, as offered in Dougherty et al. (1989) and recommended by Thoma et al. (2010), could be offered during the measurements to ease appointment arrangement. In a full RCT, it might be of advantage to organise home visits for the perineometry measurements, too. This would also resolve the argument of one potential participant who declined participation because she did not want to visit the hospital (where the measurements took place) with the baby.

Feasibility criterion 2 on the rate of participants having attended the initial pelvic floor muscle measurement within 3 weeks of consenting to take part in the trial could be fulfilled (92.9%, 95% exact CI [82.7, 98.0] vs the expected ≥ 90%). Criterion 5, the rate of participants having the final data collection within 2 weeks of ending the intervention, asked for ≥ 80% but only got 66.7%, with the 95% CI upper limit reaching 79.3%. However, it can be considered fulfilled within 3 weeks (rate 79.6%, upper 95% CI limit 90.4%). Reasons for the low rate at this criterion probably are that there were two appointments to attend within the given time– the final trial visit and the perineometry appointment–and the difficulty to find a convenient time for the perineometry appointment.

Of the suggestions to enhance adherence to measurement visits by Hulley et al. (2013), clarifying visit content, avoiding waiting time, and reimbursing travel costs were applied in this trial; notifying the participant shortly before the visit was introduced during the trial and should be kept in a full RCT. Offering the measurement appointments at convenient times of the day, especially mornings, should be enhanced in a full trial, and the two final appointments (visit and perineometry) might perhaps also be combined. The other reason for late measurement that can be influenced by the trial team is not to include potential participants before planned absences.

Data collection process in detail

For the pelvic floor questionnaire, the application process ran smoothly, although correct completion checking had to be introduced. The difficulties with Question 1 (participant reported muscle strength) of women who reported to have had a stronger pelvic floor after than before birth means that this question is not best suited for such women. An alternative would be to reformulate the question to “compare now with before the recent pregnancy”.

This however might, on the whole, be more difficult to answer because of the longer time gap to the reference value; also, as it is possible to enter values over 100%, the participants concerned have a response option. As only three (initial questionnaire) and four (final questionnaire) participants (one on both occasions) were concerned, and discussion only arose with two of them, changing question wording might be exaggerated; the more as it is not known, without testing its feasibility anew, whether the reformulated question would work. Therefore, it is suggested to leave Question 1 as it is but, if needed, to clarify the answering options with the participant on site.

Only one participant not feeling able to rate her pelvic floor muscle strength (at the final measurement) represents a lower rate than in Dietz et al. (2012), in whose study 32 of 513 participants were not able to rate their pelvic floor muscle strength. The reason here (participant may have been unsettled by the fact that she could not hold the vaginal ball inside) might differ from the (not stated) reasons in Dietz et al. (2012). It is possible that Dietz et al. (2012) left the participants alone to fill out the questionnaire, whereas in this trial CO was present during data collection and questions could be clarified. It is also unclear whether Dietz et al. (2012) provided a response scale as was the case in this trial; instead they may just have asked for a percentage number.

One participant suggested to enquire subjective pelvic floor perception more thoroughly. However, with respect to pelvic floor muscle strength, this suggestion is difficult to realise as there is only one more question in Thibault-Gagnon's (2014) pelvic floor and birth questionnaire referring to the aspect of muscle strength, and no other subjective pelvic floor muscle strength measure is available.

The collection process for adherence data was complicated by lost adherence sheets. Participants' feedback suggested to provide electronic versions of adherence charts as an alternative. According to Deshpande et al. (2011), advantages of an electronic as compared to a paper-based PRO collection tool are avoiding data entry errors and reducing the amount of missing information; barriers to the use of electronic PRO data collection might be increased expense, time needed for participant training, and infrastructure needed at the study site.

Withholding assignment information from blinded data collectors is a weak strategy (Meinert, 2012), which was confirmed by allocation having been revealed to the perineometry assessors in two cases. Therefore, a more thorough instruction to participants regarding the blinding of outcome assessors is recommended for a full trial. This could be done by including a repeat instruction in the routine text message sent before each measurement

appointment. To remind participants at the beginning of every meeting not to mention their allocation could be added to the measurement standard.

Some details mentioned by participants point to the importance of better assuring standardisation of perineometry measurement. A challenge was the accompanying baby/children needing attention. The distraction during the measurement, or the baby lying on the mother's chest, may have diminished measurement accuracy, and the babysitting service suggested to ease appointment scheduling would therefore also be helpful to ensure accurate measurement. Two participants doubted the reliability of their own measurement, and endurance measurement seems to have evoked the impression of limited trustworthiness in two of the three assessors when they were not able to explain the origin of the exceptionally high values. Acccurate endurance values might be more likely with better instructed participants, more experienced (or better trained) assessors and closer result monitoring by the PI. When checking the participants' ability to contract their pelvic floor muscles at the information/consent/initial visit, as was recommended on page 190, the measurement might be better performed according to the necessary standard.

Two participants mentioned their difficulty to have the perineometry measurement in a supine position. Although other body positions are possible for this measurement (Bø and Finckenhagen, 2003), it did not seem that they meant to suggest another measurement position. Therefore, and as supine is the position with the highest acceptance (Bø and Finckenhagen, 2003), this measurement position should be kept. Rare uneasiness about a vaginal probe entering the vagina (spontaneously expressed by few participants) has also been described by Cammu et al. (2004). In contrast to the respective PPI consideration and participants' feedback in Cammu et al. (2004), none of the nine participants for whom this was the case mentioned measurement by a male professional as a problem.

Participants were interested in and informed about their measurement values after having completed the final pelvic floor questionnaire. Although they would have liked an interpretation of their values ("Is my pelvic floor strong enough?"), this was difficult as no classification system of values was available at this time; women were however informed about the approximate position of their individual values within the pool of measured values. In the meantime, Angelo et al. (2017) have developed a classification scheme for Peritron values from very weak to strong according to the Oxford pelvic floor grading system[102] which could be used for this purpose. However, there is no specific voluntary contraction strength

[102] These authors propose to classify values below 7.5 cm H_2O as equivalent to level 0 on the Oxford scale, values from 7.5 to 14.5 cm H_2O as very weak pressure, from 14.6 to 26.5 cm H_2O as weak pressure, from 26.6 to 41.5 cm H_2O as moderate pressure, from 41.6 to 60.5 cm H_2O as good pressure, and above 60.6 cm H_2O as strong pressure. Their level 0 is at odds with the observation in the present trial that participants could have palpable contractions but perineometry values below 7.5.

that needs to be attained. The point is to be symptom free, and as the correlation between pelvic floor muscle performance and symptoms is not straightforward (Theofrastous et al., 2002), the aim is only a gain in strength.

Interviewing about adverse events will remain essential in a full RCT. In this feasibility trial, the interview probing process for adverse events was suboptimal because “muscle soreness–vaginal discharge–infection” were listed on the form to aid remembering yet did not have a yes/no option prepared per item. Although it was enquired with care, this suboptimal form design may have introduced bias towards missing harms. To enhance harms screening, a more comprehensive checklist with a yes/no option for each item should be designed for a full trial. Further, monitoring recurrence of adverse events should be elaborated.

The interviews served the feasibility trial purpose. In a future RCT, the preintervention interviews therefore are not needed any more. With the potential modification of the experimental intervention to enhance ball weight however, the postintervention questions on participants’ clinical experiences should be kept in interview form. It is also recommended to again survey participants’ opinion on the interventions in online form as an anonymous survey offers the opportunity for honest feedback. Participants also should be asked again about their experiences with and opinion on the trial but this should comprise less detail and be left mainly to the online survey. Although it had been intended to pilot the survey data collection tools for the main trial, the postintervention interview schedule in its current form is not applicable to a full trial. With the experience gained, the schedule/questions can be refined and adapted to the new situation. Likewise, the content of the online survey needs to be reconsidered and the questions need to be adapted.

Interventions

In the following three sections, the results about the trial interventions–first intervention uptake and then their feasibility and adherence to them–are discussed.

Intervention uptake

The date of intervention start was collected in two ways: noted by the researcher at the first adverse events call which took place approximately four days after initial perineometry measurement, and by the participant on her adherence sheet. However, these dates were only congruent in 39 cases. In eight cases they were not congruent, and in eight cases date congruence was not applicable or could not be examined (e.g. never started intervention, lost adherence sheet). The incongruence found was between one and 15 days (mode 1, median 2, with one unclear), and in these divergent cases, the date which most likely seemed to be correct was chosen as the valid one.

Feasibility criterion 3 on intervention start was not fulfilled. Instead of ≥ 90%, only 81.8% (95% CI [71.6, 92.0]) of the participants started the intervention within 4 days of the initial pelvic floor muscle measurement; however, the CI includes the aimed for value, and the aim to start within 4 days was set arbitrarily (as a timely start is desirable) and can be questioned. After 7 days, 89.1% (95% CI [80.9, 97.3]) of the participants had started the intervention. Considering the reasons given for not starting sooner, the only feasible enhancements in a full RCT could be to not let women enter the trial before planned absences or to inform participants more clearly to start the intervention within 4 days of the initial measurement. Towards a more exact documentation of intervention start in a full trial, electronic versions of adherence charts and more explicit clarification at the corresponding phone call might enhance the situation.

Feasibility of interventions

Although for some women it is physically not possible to use vaginal cones (Bø, 1995b, Herbison and Dean, 2013, Bø, 2015b), it was possible for all participants in this trial to use the vaginal ball. Ball use was experienced positively (occasionally even as fun) by about two thirds of the participants. Contrary to Bø et al. (1999, p. 491), participants in this trial did not report “motivation problems and trouble” in using the device (discomfort or pain at ball use will be detailed in section 8.3 on harms). Ball use best fit into regular daily life at home, confirming Test-Club-Bericht (2013). A few participants used the ball outside of the house which posed practical difficulties because of the restriction to 30 minutes and the question of where to take out and clean the ball. Adhering to ball use for exactly 30 minutes however also was an issue at home. Some participants set a timer (one of them suggested to provide a timer to participants in a full trial), others used the ball for longer than 30 minutes. Despite this, the instruction to wear the ball for longer than 30 minutes, as suggested by a participant, does not seem realisable as 30 minutes is already at the outer limit of the duration of device use in cone trials (Herbison and Dean, 2013). The need to remember doing the intervention was named as a disadvantage, as was being bound timewise or to the house. As enhancing ball weight is an option for a full RCT, feasibility of ball use might be different with this modified intervention application.

Correct placement of the ball as a learning process was addressed by five participants. At some point having lost the ball was reported by seven, confirming Test-Club-Bericht (2013) which reported one participant with initial difficulties to hold the ball in the vagina. All except one participant could find a reason for this and solve the situation. Similar to the participants in Jonasson et al.’s (1989) study who found cone handling simple, all participants in this trial found ball handling simple; only cleaning issues were named which are rather specific to the ball model used. Also specific to the ball model, the retraction cord was described as too

short and of suboptimal material (as in Test-Club-Bericht (2013)). As ball cleaning to medical hygiene standards after vulvovaginal symptoms would have been too difficult to organise, and to avoid hygiene problems by incorrect disinfection, participants instead were provided with a new ball. For a full RCT, a solution to this hygiene question needs to be worked out. Antibacterial soap samples for travelling, a participant suggestion, might be organised in a full trial.

About a third of the participants performed pelvic floor muscle exercises additionally to ball use. However, only about half of them did so with the frequency recommended for evidence-based pelvic floor muscle training; also, participants in this group were not explicitly introduced to the concept of exercise science for pelvic floor muscle training but referred to the exercises recommended by their caregivers, which makes it is likely that they did not perform exercises according to exercise science recommendations. Although it was not clarified in the written instruction for vibrating vaginal ball use that being in the experimental group should not discourage from doing routinely recommended pelvic floor muscle exercises, this was explained orally at trial entry to every participant in this group. However, this oral instruction may not have been noticed by every participant concerned. This became evident when one participant asked during the final interview whether she had been allowed to exercise. Two of the participants who did not perform any exercises emphasised that they had taken advantage of their “laziness” (ID 22) and would not have done the exercises without trial participation either.

In spite of the difficulties named (including discomfort/pain issues of 11 participants, see section 8.3), most of the participants could imagine continuing ball use. This is at odds with other studies. In Fischer and Baessler (1996)/Fischer et al. (1996), a nonrandomised trial on postpartum cone use by mostly continent women, slightly under half of the participants had good or very good motivation to continue cone use, and motivation was low in a third (including those not having shown up for final assessment). In Williams et al. (2006), a study using cones in urinary incontinent women, only 56% in the cone group were motivated “a lot” after four weeks. In Cammu and Van Nylen (1998), none of the participants wanted to continue cone use after the study period of 12 weeks. However, 80% of the testers in a consumer journal test report would recommend the balls used in this trial to other women (Test-Club-Bericht, 2013).

The pelvic floor muscle training experience in the comparison group was rated positively by almost half of the participants in the interview and by all online. Training best fit into daily life at home, during baby feeding or in bed in the evening; on holiday, one participant had more time to train whereas others had less. The need to remember doing the training was not named as a disadvantage of the intervention but came up with the questions on difficult

aspects of and barriers to the training. Participants, although not all, seem to have varied their training positions (not enquired systematically). Although a vaginal check of correct contraction was performed in this feasibility trial in addition to explaining the exercises, one participant fed back that she would have liked better training instruction; however, as aiming to keep to standard care, a better explanation of training might, depending on the setting, surpass usual care in a full RCT. Almost all participants in this group could imagine continuing pelvic floor muscle training. As further studies are required to find the optimal training dose (Mørkved and Bø, 2015), the exercise regimen in future postpartum pelvic floor rehabilitation guidelines and in a future trial may change, and one participant's suggestion to prescribe fewer exercises might be realised.

Adherence

Research participation and adherence measurement lead towards better adherence[103] (Horne et al., 2005). Indeed, in Cammu and Van Nylen (1998), commitment to take part in the study was the reason why most of the participants continued using cones; this was also expressed colloquially by a participant in this trial: "If I had not been a trial participant, I would have finished sooner". Adherence in this feasibility trial was potentially enhanced by drawing participants' attention to their intervention by adherence diaries and adverse events phone calls (as in other pelvic floor trials described in Woodley et al. (2017)); one participant indeed mentioned having used the adherence sheet as a reminder by placing it where she could see it.

Despite this, Feasibility criterion 4b on adherence was not fulfilled. Depending on the calculation method, only 47.2% (95% CI [33.8, 60.6]) minimum or 60.4% (95% [CI 47.2, 73.6]) maximum of participants had adequately adhered to the intervention, with roughly similar rates in both groups. This means that even as trial participants and with wider than planned adherence criteria, 40% of women were not adherent. More participants (6 vs 5) spoke of having forgotten the exercises than the balls–this could suggest the ball has a reminder function as may have been the case in the trial by Porta-Roda et al. (2015). As participants in this trial were explained its feasibility nature and that adherence was an outcome measure to determine whether the trial was feasible, they may have been relaxed in terms of reaching adherence (as one participant mentioned).

Of strategies suggested to enhance adherence (Laycock, 2008, Hulley et al., 2013), some have already been used in this feasibility trial: choosing participants who are likely to adhere, checking adherence, and encouraging participants by phone calls. Further, pelvic floor

[103] A reactive effect to being studied (e.g. increased adherence) is called the Hawthorne effect (Burns, 2000, LoBiondo-Wood and Haber, 2013a).

training programmes supervised by a health professional are more useful than simple verbal advice or an unsupervised programme (Bump et al., 1991, Bø, 1995a, Alewijnse et al., 2007, Hay-Smith et al., 2011). In this feasibility trial, contacting participants by adverse events calls can be seen as a form of supervision, and one participant mentioned the phone calls as important to keep her going and feel valued.

Adherence was measured indirectly by participant self-report and may therefore be subject to social desirability bias with participants documenting a level of adherence which they think is desired by the researcher. The adherence jotted down does not necessarily represent the "real" adherence, and the results therefore should be seen as "indicator of adherence" (Horne et al., 2005, p. 44). Horne et al. (2005) assume that disclosure of nonadherence is more truthful than that of adherence, which opens the potential for adherence in this trial to be even lower than calculated. The adherence rate may also be subject to recall bias by the fact that a few participants did not note adherence immediately after the intervention or by replacement forms having been completed at least partly in retrospect.

An adherence data collection/analysis issue which arose during the final discussions with participants concerns the "real walking time". After a participant had mentioned to have substracted sitting time during ball use, this was enquired more thoroughly in the following participants. As far as enquired, the tendency was that most participants had substracted sitting time during ball use and noted net minutes. It was therefore assumed that the documented 30 minutes represent walking time; however, this issue needs to be given attention in a full RCT. This also means that more participants than those who had noted or mentioned this may have left the ball in the vagina for more than 30 minutes. One participant suggested to tick adherence on the form instead of noting minutes of ball use but this is not possible as shorter and longer than intended ball use need to be documented. Similarly, a few participants sometimes performed more than the requested exercises per day; the participant who often performed 20-30 (a few times even 100) short contractions seemed not to have understood the training principles which points to the need for better explanation of these principles.

Identified barriers to adherence are similar to some of those found in studies on pelvic floor physical therapy (Alewijnse et al., 2007): forgetting to exercise, difficulty to integrate exercising into daily life, stressful situations, lack of time/motivation/discipline, and perception of (no) symptoms. Specifically to pelvic floor muscle exercises following a perineal tear, Gillard and Shamley (2010) also identified being occupied with or distraction by baby, both also expressed in this trial as "finding time with the child[ren]" or initiated but not completed exercise blocks. For the postpartum period, Fine et al. (2007) further identified that exercises can hurt (in their study, exercises had hurt for 6% of the women who never performed pelvic

floor muscle exercises and for 2% for those who had ceased to do so by six months post partum). As in other studies (Mason et al., 2001a, Alewijnse et al., 2007), perception of (no) symptoms may have influenced motivation and adherence, as 34 of the 54 women who completed the trial did not have symptoms. Cone studies did not investigate adherence (Wilson and Herbison, 1998, Herbison and Dean, 2013).

The group preference seems to have played a role for facilitating adherence as indicated by the slightly higher motivation after allocation to the preferred group and the somewhat higher adherence rate of women in the experimental group (the preferred group). In one experimental group participant's words: "In the exercise group it would have cost me even more of an effort [to adhere]" (ID 51). However, direct comparison is difficult because of the different ways of measuring adherence in the groups.

Other forms of adherence calculation, e.g. by ordinal or dichotomous categories, may have led to a different result. As different methods of adherence measurement and/or reporting are used in different studies (Mason et al., 2001a, Woodley et al., 2017), comparison across studies is difficult. Depending on the method of measurement, adherence in other studies on pelvic floor muscle training after childbirth to prevent urinary incontinence and including both women with and without urinary incontinence was between 16.5% and 100% (Mørkved and Bø, 1997, Chiarelli and Cockburn, 2002, Hilde et al., 2013a). However, Chiarelli and Cockburn's (2002) trial studied a pelvic floor rehabilitation programme with adherence incentives; these authors, like Mørkved and Bø (1997), only used an 8-week intervention period, but this trial was of 12 weeks duration.

As conclusion about the interventions it can be said that, despite some difficult aspects, ball use was feasible as a trial intervention, as was pelvic floor muscle training. In a future trial, the difficult aspects can be communicated to potential participants when discussing participation. Although it is not clear how accurate the adherence estimate is, measures to enhance adherence are needed (and laid out in Chapter 10 on future research). The need for a differentiation between level of adherence measured versus level of adherence needed for an effect is elaborated in the effect results section.

10.2.2 Trial management and resources

Human, organisational and data management findings could be compiled. As this trial's human management findings, such as the challenge to find assessors and organise a rota for the perineometry measurements, are specific to the PhD circumstances of the project, the information gained might not be fully applicable to other research settings. Due to the voluntary nature of the assessors' contribution, perineometry appointments in this feasibility trial could only be offered once or twice a week. From the experiences gained it seems

advisable however to offer appointments more often, which could be arranged with paid assessors.

Similarly, with a budget available, a more friendly perineometry venue with a computer to access the data collection form (which would enhance data protection) and a locker in the measurement room to store the Peritron and transport tickets could be afforded. Participants in this feasibility trial appreciated home visits for the research meetings which could be offered because of the PhD nature of the project. In a “real world” research setting, this might however not be affordable even with a budget available.

A challenge with respect to ethics committees might, depending on the country, be the intimate nature of the perineometry measurement. Although a vaginal assessment is intimate, participants’ and assessors’ feedback seems to indicate that, similar to Chiarelli et al. (2003), participants were not or only very rarely embarrassed by the situation. Suggestions for trial management enhancements are to give perineometry assessors an understanding of the research ethics involved and to perform risk assessment on their behalf. With the help of this feasibility trial experience, document numbering, file naming and variable coding can be devised for a full RCT. Further, it is suggested to use the statistical computer programme R instead of SPSS.

As in Tickle-Degnen’s (2013) feasibility trial, the most systematic management assessment in this feasibility trial involved compliance with ethics and data management. Other aspects of management assessment were completed more informally and on the basis of research experience without a guiding framework (as none available). The author has taken great care to capture all relevant topics comprehensively, nevertheless mistakes or omissions may have crept in.

Resources needed for this feasibility trial were compiled for the categories facilities, staff, materials and transport. The company producing the vaginal ball used did not fund the balls and probably will not do so for a full trial. A future funding option could be to use balls of another company willing to support this research; however, this would partly nullify this trial’s feasibility results as the balls have been chosen by PPI participants and shown feasible in use during the trial which might not be the case with different balls; it might also nullify the preliminary clinical results. Unplanned expenses were incurred by necessary replacement balls and waste visits. In a full trial, additional costs must be planned for balls of increasing weight (three per participant when using Laselle Weighted Exercisers), and there should be a budget available to pay (potential) participants’ refreshments for meetings at public venues, or for staff meetings. Potential to lower trial costs is twofold: Refunding transport costs could be minimised by refunding only when a participant indeed incurs costs, and if the research

project was affiliated to a hospital, an appropriate smaller amount of lubricant could be bottled in the hospital pharmacy at probably lower cost than buying it commercially.

The resources calculation of this feasibility trial, complemented with Tickle-Degnen's (2013), Thabane et al.'s (2010), and Orsmond and Cohn's (2015) resources questions, partly reproduced in Appendix CC, can serve as the basis for a budget calculation for a full RCT. This budget calculation has to be adapted to the future intervention (which might encompass enhancing ball weight), the number of participants, and the setting and circumstances at the potential future trial site; the precise budget needs to be calculated according to the respective local costs.

10.2.3 Clinical findings

The discussion of the clinical findings first covers the participants' clinical experience with the experimental intervention, and after this data from PROMs, perineometry and adverse events screening.

Participants' clinical experience with vibrating vaginal balls

Plevnik (1985) purported the sensation of losing the cone as the method's working mechanism. However, Bø (1995b) argued that it was not verified whether all women have the sensation of the cone slipping out. As recommended by her, this argument was followed up in this trial. The scientific enquiry shows that, with a ball of 28 g, half of the participants did not feel the sensation of the ball wanting to slip out of the vagina and only a quarter felt this at least sometimes. In fact, about a third of the participants reported that they had not noticed the ball during use. However, a ball of 28 g might not represent the heaviest weight participants might be able to hold and with which they might have experienced this sensation. Another reason for not sensing the ball might be that it was placed too high in the vagina.

Pelvic floor reactions purported for cone use are reflectory and voluntary contractions (Plevnik, 1985, Peattie et al., 1988, Bø, 1995b, Bø, 2015b). This was confirmed by only four and eight participants respectively who felt automatic enduring pelvic floor contractions or performed voluntary contractions during ball use; more than half of the participants did not notice a pelvic floor reaction. The described effect of enhanced blood flow and sensitivity (Schildbach, 2005) might be confirmed by the few participants who sensed pelvic floor "activity".

Considering Bø's (1995b, 2015b) argument that putting a weight on the pelvic floor muscles could further fatigue instead of strengthen the pelvic floor muscles, it is possible that some of the participants might be classified as having experienced pelvic floor fatigue. Three participants mentioned to have felt pressure from the inserted ball; the participant who

repeatedly lost the ball might have experienced this because of easily fatigued pelvic floor muscles or because of too enlarged a hiatus (Shek and Dietz, 2009). A quarter of the participants had the ball fall out at some point, whereby participants attributed this to the use of lubricant, day of the menstrual cycle, or strain on the pelvic floor. The test club article (Test-Club-Bericht, 2013) also reported one participant for whom it was difficult to hold the ball initially. Two facts mentioned on vaginal cone use by Stewart (2006) were confirmed by two trial participants: One is that cyclical variation in vaginal secretions affects device retention, the other is that due to fatigue of the pelvic floor muscles during the day or by preceding sports activity, the weight can be better retained in the morning than in the evening. This latter observation was also reported by Cammu and Van Nylen (1998).

The vibrations were felt by most participants but often only with stronger movements, and all the time only by four; three participants indeed mentioned to have set a timer to remind them of the end of the 30 minutes of ball use, and one suggested to provide a timer to participants in a full RCT. The vibrations might be more noticeable with products from other companies, heavier ball weight, double balls (as one participant reported from her earlier experience), or with an electronically vibrating device (as e.g. by Amorelie (no date)). However, feeling the vibrations might not be necessary and the intervention might be as effective (or not) if women do not feel the vibrations. Although vibrations are said to contribute to the balls' effect (Schildbach, 2005, ELANEE, 2017b, FUN FACTORY, no date-d, a, MAPA GmbH, no date-b, medesign I.C. GmbH, no date-b), only one of the named sources is implicitly purporting that women can *feel* the vibrations, suggesting that it would depend on the state of the pelvic floor whether women can feel ball weight and vibration (FUN FACTORY, no date-a). Not all women feeling a pelvic floor reaction was also the case in a feasibility research project on special shoes to strengthen the pelvic floor (ZHAW Zürcher Hochschule für Angewandte Wissenschaften, 2016, no date, publication forthcoming).

Pelvic floor muscle performance

The following two sections comprise the discussions with regard to issues specific to PRO and perineometry results respectively, whereas topics applying to both these methods of pelvic floor muscle performance measurement are considered in the third section.

Discussion specific to participant reported outcomes

The PROs comprise the quantitative pelvic floor questionnaire results and the qualitative results from the open-ended interview questions and are now discussed in this order.

Quantitative PRO results

Considering the decreased prevalence of urinary incontinence after birth compared to pregnancy (see section 2.3.3), it is surprising that only six participants rated their pelvic floor

strength post partum higher than before birth. This is however proportionally more than the four of 481 participants who reported increased strength post partum compared to before birth in Dietz et al. (2012). Within each intervention group, participant reported pelvic floor muscle strength rose from pre- to postintervention; with -5.1% (95% CI [-18.0, 7.8]), a change difference between the groups with a higher rise in the comparison group was found. However, the broad CI points to great uncertainty.

This result for participant reported pelvic floor muscle strength is supported by the result for postintervention Question 2, where participants rated the pelvic floor more often as "enhanced" in the comparison group. Again, the RR of 1.23 shows a wide 95% CI [0.95, 1.60] and thus uncertainty. Accordingly, the descriptive symptom values shrank more in the comparison group for all symptoms (except for stool incontinence which was not comparable). The bothersomeness results might indicate that the participants who still are symptomatic at the end of the intervention period might be more bothered by this than at trial entry–they might have expected symptoms initially postnatally but be (more) bothered by them if they have not resolved after a certain time; the participants who were bothered "not at all" at trial start may have been cured at trial end. It is remarkable how many participants reported pelvic floor symptoms and were bothered by them although they had confirmed that they had been discharged from postnatal care as "everything ok".

Further supporting the above results, the ICIQ-UI SF sum score mean reduction and the decrease of the urinary incontinence rate was larger in the comparison than in the experimental group. Besides stress, urge, and mixed incontinence, other forms of urinary incontinence were indicated on the ICIQ-UI SF. Although pelvic floor muscle training might not be the best intervention choice for leaking after urinating or for no obvious reason, three of the four participants initially concerned indicated no symptoms at trial end, while the one who did was in the experimental group (originally two in each group).

When completing the pelvic floor questionnaire at the final study visit, participants may have remembered their preintervention PRO responses. However, when asked, only few participants confirmed this (mostly of 100% ratings). Also, a placebo effect[104] in the self-improvement rating cannot be excluded: in the intervention group because of a belief in the ball's effectiveness, in the comparison group because of a sense of activity and self-control (as expressed by ID 42: "I heard from ball group participants [her cousin and others when waiting for measurement] that using the ball was less strenuous than doing exercises"). At the final interview, it turned out that a few participants had taken pictures of their

[104] "The beneficial effect in a patient following a particular treatment that arises from the patient's expectations concerning the treatment rather than from the treatment itself" ('Placebo effect', 2007, https://medical-dictionary.thefreedictionary.com/placebo+effect).

perineometry results. Per se desirable as participants should have the information about their outcomes, this may have biased the results in that participants may have rated their final PROs differently after having seen their perineometry values' change. The answers about an enhanced pelvic floor (postintervention Question 2) and the four symptom and bothersomeness questions were dichotomised, and the method of dichotomisation may have influenced the results.

A comparison of the present quantitative PRO results to previous work is difficult as there was no study found with the same or a similar intervention post partum using these measures. The results for Question 1 can only be compared to those of the observational study by Dietz et al. (2012) where the question originated from. The average strength in Dietz et al. (2012) was 89.1% (range 0-110) which is higher than the 74.6% pre- and 85.9% postintervention mean in the present trial. The reason for the lower preintervention mean in this trial might lie in the fact that Dietz et al.'s (2012) data were collected at three to six months post partum whereas in this trial they were collected between six and 24 weeks after birth. Another difference is that Dietz et al. (2012) asked for contraction strength estimate; they could do this because all participants had received pelvic floor muscle exercise teaching by professionals and had been encouraged to perform the exercises ante- and postnatally. In this feasibility trial, participants were (implicitly) asked for pelvic floor muscle strength independent of contraction as experience in correct pelvic floor muscle exercises could not be expected. Therefore, although the question has shown reliability in its original study, comparing pelvic floor muscle strength irrespective of contraction in this trial might invalidate its test-retest repeatability result (Dietz et al., 2012).

Wilson and Herbison (1998) reported participants' postintervention rating of "adequate vaginal tone" with respect to sexuality, which may be compared to this trial's symptom question "Do you think that your vagina is too loose or lax?". They did not find a statistically significant difference between the groups for the category "less than adequate"; however, the trial's limitations (as discussed in section 3.4.2) need to be taken into account. Wilson and Herbison (1998) also collected self-reported urinary incontinence rates one year after delivery. This was nearly the same in the cone and (enforced) exercise groups, and higher in the control (routine care) group.

Qualitative PRO results

A wording weakness was identified with the initial interview question formulation on pelvic floor changes *since* birth. It had been intended to enquire changes *by* birth; as one participant answered the question by telling about her pelvic floor enhancement since birth, it was modified to clearly express changes *by* birth. Also, this question did not include

pregnancy as contributing to pelvic floor changes, which mirrors the respective weakness of Question 1 on participants' self-rating of pelvic floor muscle strength.

As answers to the open-ended interview question, pelvic floor changes after the intervention were reported more frequently in the comparison group, with the pelvic floor more often being enhanced and symptoms more often being reduced. One participant in the experimental group described her pelvic floor as worse than before the intervention. The same tendency in favour of the comparison group can be seen in the qualitative online survey results in Table 11 (intervention feasibility section 6.6.1). The pelvic floor muscle performance results for the ball group are consistent with those of Fischer et al.'s (1996) nonrandomised postpartum cone group, in which participants also stated that their pelvic floor had become firmer and the vagina tighter (no comparison with exercise group).

Enhancing awareness of pelvic floor muscles, a basic component of pelvic floor muscle training (Baessler and Bell, 2008), is a function ascribed to vaginal cones (Thakar and Stanton, 2000, Rochera et al., 2017, LELO, no date, MAPA GmbH, no date-a). This function was confirmed for ball use by four participants in this trial, similar to the participants in Fischer et al.'s (1996) cone study and in the ball test reported in Test-Club-Bericht (2013). Enhanced pelvic floor muscle awareness can support the device function of helping women to learn to train their pelvic floor, as mentioned in medesign I.C. GmbH (2018c) or Fischer et al. (1996). If of interest, a full RCT could include pelvic floor muscle awareness and enhanced ability to contract as secondary outcomes.

Nearly a third of the participants had not or only to a minor degree re-established their (respective) sexual relations at the time of the final interview and thus was not able to comment on the interventions' influence on sexual sensations. The necessity of explicit probing about the resumption of respective sexual relations by a screening question before asking about changes in sexuality was realised after the first few cases and then introduced with an informal screening question. There therefore are a few ambiguous answers where it later was not clear whether the noted "no" meant "no change experienced with sexuality" or "no change experienced as no sexuality". This screening question must be added on the interview schedule for a full RCT (e.g. as "Have you been sexually (vaginally) active since the birth of your baby?").

The described changes were enhanced sensation or muscle tone, a heightened pelvic floor consciousness, and encouragement to more intense sexuality (ball group). In accordance with the other PRO results, more participants in the comparison group described changes with respect to sexual sensations. In Wilson and Herbison's (1998) postpartum trial, participants were asked about sexual satisfaction one year post partum. For the outcome

"vaginal feelings", the result was similar between the intervention (comprising intensive exercising, cones, or both) and control group (comprising standard care; cave: trial limitations, see page 66). In Fischer et al.'s (1996) cone study, postpartum participants reported more sexual pleasure after cone use (no comparison with exercise group). As with any interview data collection, missed information and interviewer bias may have happened.

Discussion specific to perineometry

After one month of exercise, a pelvic floor muscle strength gain of 100% could be found (Bø et al., 1990a, Bø and Aschehoug, 2015). According to sports science, trained individuals have slower rates of strength improvement than untrained individuals (Kraemer and Ratamess, 2004, American College of Sports Medicine, 2009, Bø and Aschehoug, 2015); for the pelvic floor, it was shown by Theofrastous et al. (2002) that women with low initial contraction strength achieved a greater reduction in urinary incontinence symptoms than women with higher initial contraction strength. The maximum strength values found in this feasibility trial behave according to the described patterns, their strength rise was between –42% and 268% (mean 40%), with 13.0% of the variance accounted for by initial value.

The fact that postpartum women's perineometry values can shrink over a trial period was confirmed by a fellow researcher (Leister, 2016)[105], in whose dataset between 17.4% and 50.0% of strength values decreased (in groups with sample sizes of 15-23). In an attempt to investigate regression of unusually high preintervention values to the mean as a possible explanation for a decrease of maximum strength values (Petrie and Sabin, 2009), it was found that only three of 10 preintervention values above one *SD* (14.6 for a mean of 23.0) had decreased. Some of the participants with reduced postintervention values were asked for a possible explanation from their point of view. Their ideas to explain the decrease of their maximum squeeze pressure values are collected in Box 9. Some of the explanations can be shared from a researcher's point of view, e.g. that muscle strength also depends on psychological factors (DiNubile, 1991), or that psychological stress could have been induced by the assessment itself (Dietz and Shek, 2008a).

[105] At that time postdoctoral research fellow at the School of Nursing, University of São Paulo, Brazil.

Box 9 Some participants' ideas to explain the decline from initial to final perineometry values for maximum strength squeeze pressure

- Different instruction?–"you know already" (ID 16)
- Activated gluteal muscles at initial measurement without assessor noticing it
- Felt differently at final measurement, stress
- With exercises also, it was sometimes easier, sometimes more difficult
- No measurement explanation at final measurement because of lack of time (participant was in a hurry)
- More embarrassed at final measurement
- Exercised pelvic floor less but did more sports
- Baby on chest at initial measurement (subjectively enhanced)

As perineometry resting pressure means were expected to rise (Griffin et al., 1994, Hilde et al., 2013a), it is unclear why they diminished in both groups, with the smaller decrease in the experimental group. The maximum and mean values for contraction strength rose more in the experimental group, while the endurance pressure rose more in the comparison group. However, all effect sizes are small, and the broad 95% CIs point to high uncertainty; all change score difference CIs include no effect, and change scores might as well, as can be seen in Figure 26, run in the opposite direction. As all perineometry values declined in the participant who could not hold the ball in the vagina, her exclusion in an explanatory analysis would have led to less conservative values.

There is no standard duration for a perineometry strength contraction. In this trial, the duration was not specified and was shorter than in studies on perineometry reliability and in observational studies (see Appendix AA). Since in many studies the duration was longer than in this trial, a modification to consider for a full trial is to ask participants for a holding time of 2-3 to 5 seconds to align with other work in the area. The duration of the interval between contractions may also be a confounding variable, and the short break interval of 10 seconds (which however is in line with other studies (Hundley et al., 2005, Friedman et al., 2012)) may potentially have had a negative influence on the accuracy and reliability of the perineometry measurements. A longer break between measurements may minimise the effects of fatigue on the contractions. When endurance was, in rare cases, measured twice (see section 6.5.3), the assessors did not extend this break which might further contribute to bias.

As detailed in Appendix AA, reliability of the Peritron is fairly high for resting pressure and high for squeeze pressure strength (correlation coefficient values, depending on inter- or intrarater testing and body position, between 0.74 and 0.97). Reliability has not been established for squeeze pressure endurance as operationalised in this trial, as the three respective reliability studies operationalised endurance differently. Every attempt was made

to ensure standardised measurement; however, perineometry requires skills (Bø, 2015c), and the assessors at trial start were “perineometry novices”. In the cases with uncertainty about which of the communicated values to accept as the accurate one, the available values were pondered and the most likely one chosen for analysis. In spite of doubts on endurance reliability, the endurance values as a summary measure raised similarly in both groups. Three outcome assessors taking measurements enlarged observer and thus data variability[106] and increased measurement error (Lenth, 2001, Petrie and Sabin, 2009). This was confirmed by spontaneous comments from participants about differences in the instruction (although instructions had been written down in detail in the measurement standard).

Two participants mentioned that time of the day had made a difference for them for ball use, which was confirmed in the literature (Stewart, 2006). Although Dougherty et al. (1991) concluded that time of the day does not influence perineometry results, a strategy for enhancing their reliability in a future full RCT might be to standardise or record time of the day at measurement (as suggested by Messelink et al. (2005)). In terms of measurement physiology, a newly developed portable dynamometer not yet available at the start of this trial might be of interest (Univalor, no date). A dynamometer’s advantage against a perineometer is that it is not influenced by abdominal pressure rises but exclusively measures vaginal squeeze pressure (Dumoulin, 2004, Hanzal et al., 2015, Univalor, no date).

Vaginal pressure results cannot be compared between studies that used different models of vaginal probes with different diameters (Bø et al., 2005, Barbosa et al., 2009), and a comparison of the present results to previous studies also needs to take into account the feasibility nature of the present data. Nevertheless, the high variability of the vaginal pressure measurement results of the present trial is consistent with that found in earlier work (for a collection of postpartum perineometry squeeze pressure strength value variability in other studies see Appendix GG). Wilson and Herbison (1998), the only postpartum cone study, found a *SD* of 9.6 cm H_2O in the cone group, of 8.2 in the control group (routine pelvic floor care), and of 8.4 in the (enforced) exercise group for muscle strength, and of 7.7, 6.1, and 5.9 cm H_2O for endurance respectively (calculation by CO, Table 2). The tendency of the perineometry result for contraction strength in this trial (experimental intervention better) is in contrast to that by Wilson and Herbison (1998), where at trial end the cone group had the lowest values, followed by the control group, with the exercise group being highest. For pressure endurance, this trial’s results confirm Wilson and Herbison (1998) who found the

[106] Interobserver variation is “the amount observers vary from one another when reporting on the same material” ('Observer variation', 2008, http://medical-dictionary.thefreedictionary.com/observer+variation).

lowest values in the control group, followed by the cone group, with the exercise group displaying the highest values. However, as in this trial, the differences were small, the authors did not compare change scores but final values without consideration of initial values, and the reanalysis had limited validity (see page 66).

General discussion of pelvic floor muscle performance

Some topics concern both PRO and perineometry results. These are considered in the following.

Comparisons

Data examination at the participant level suggests that the planned outcome measures are sensitive to the effects as changes in the outcome variables occurred. Except for perineometry resting pressure, the aggregate findings are in the expected direction.

The descriptive comparison of the participant reported symptoms shows that vaginal looseness/laxity and wind symptoms declined more in the comparison group, whereas the only symptom frequency rise happened for vaginal pressure in the experimental group. Likewise, the urinary incontinence symptoms decreased more in the comparison than in the experimental group, with the two rises happening in the experimental group. The descriptive PRO results therefore seem to favour pelvic floor muscle training, which is in line with the qualitative effect results.

Comparing the direction of change between participant reported pelvic floor muscle strength and perineometry maximum strength showed that it was identical in 28 cases. In six cases, participant reported strength decreased and perineometry maximum strength increased, whereas in seven cases participant reported strength increased and perineometry maximum strength decreased. In nine cases, one measure stayed identical while the other increased, in two the participant reported pelvic floor muscle strength stayed identical while perineometry maximum strength decreased.

The calculated preliminary effect sizes (change score differences) are small and have wide CIs, including no effect and an effect in the opposite direction. The findings of the mITT and PP analyses for the two potential future primary clinical outcomes (participant reported pelvic floor muscle strength and perineometric vaginal squeeze pressure strength) were consistent. Although a PP analysis loses the randomisation effect and thus introduces attrition bias ('Intention to treat analysis and per protocol analysis: complementary information', 2012), this robustness of the results allows more confidence in the findings (Thabane et al., 2013).

The participant reported pelvic floor muscle strength favoured the comparison group. Within perineometry, vaginal resting and strength pressure favoured the experimental group,

whereas endurance favoured the comparison group. An explanation for the partly contradictory results might be the weak correlation between pelvic floor muscle strength and urinary incontinence symptoms postulated by Theofrastous et al. (2002). Part of the PRO analysis is only descriptive; a limitation of the statistical model used is its nonconsideration of initial mean differences between groups and a lack of adjustment for potentially influencing variables. Further, potential outliers of continuous measurements cannot be confirmed as such since the dataset is too small to make respective conclusions. However, the analysis is sufficient for its exploratory purpose.

Although clinical data were analysed by descriptive and comparative effect size statistics, no significance and hypothesis testing was undertaken. This is a feasibility trial limitation by its very nature as this trial design is not powered to detect a potential statistically significant effect. Although no *p* values were calculated, the effect size CIs allow a conclusion on the lack of statistical significance. However, this must not be interpreted like in a full trial as the statistically not significant results may represent an underestimation of effect by a feasibility trial's lack of power (Clark-Carter, 2003, Kraemer et al., 2006, Pocock and Stone, 2016).[107]

A lower randomisation target in the comparison group was considered appropriate as one of the research questions particularly intended to gain experience with harms associated with the experimental invervention. Similarly, the participants' experience with the experimental intervention was of more interest than that with the comparison intervention as more scientific knowledge is already available on pelvic floor muscle training by contraction exercises. Limitations of this uneven group size with respect to the clinical feasibility outcomes are a greater variability of descriptive results in the comparison group and with a larger change score CI less information on the respective values in this group. Further, by the 2:1 allocation, the effect and harm group differences are more difficult to identify and less meaningful because of the relatively broader CI in the comparison group. However, as this is a feasibility trial with preliminary clinical results without significance and hypothesis testing, this is considered a minor disadvantage.

Risk of bias

By completing the adapted version of the Cochrane Collaboration's tool for assessing risk of bias (Appendix E), a risk of bias assessment was performed for this feasibility trial's clinical

[107] To overcome the problem of unduly interpreted statistical significance, alternative options about group comparisons in preliminary studies are mentioned in the literature. These are: avoid statistical group comparisons altogether (Arnold et al., 2009), present outcomes by groups descriptively without comparing them (Bugge et al., 2013), analyse both groups in one cohort (Cook et al., 2005), report the effect size blinded by group (Arnold et al., 2009), or test the hypothesis about a future trial's primary outcome measure with a higher type I error rate to less likely miss a truly effective intervention for further testing (Schoenfeld, 1980).

results in the same way as for the literature included in this thesis' systematic review. The results table for this assessment in Appendix HH shows a low risk of selection bias, attrition bias, detection bias[108] for external measurement, and reporting bias; it shows a high risk of performance bias and detection bias regarding PROs. Forms of potential bias were also considered according to Higgins et al. (2011a), Indrayan (no date), Jadad and Enkin (2007), Kiene (2001), Sackett (1979), and Torgerson (2014). While this revealed a low risk of "other bias" for this feasibility trial in the Cochrane tool, it also led to the identification of risk of bias in domains not named in the Cochrane tool, which are collected in Box 10.

Thereby, these risks of bias can be inherent to the trial design in general, to the feasibility design in general, to this particular RCT and thus be unavoidable in a future definitive trial, or originate from this feasibility trial's design or execution. They can affect both or only one study group(s), resulting in an unclear or likely direction of risk of bias, and with a likely direction can lead to a potential over- or underestimation of an effect. Consideration of all risks of bias in Box 10 results in a judgement of high risk of bias for selection bias, and an at least medium risk of performance and detection bias.[109]

The participant excluded from the trial was also excluded from the clinical outcome analysis. This exclusion of a participant from analysis who after randomisation became ineligible is legitimate and not introducing bias when (1) the participant never received the intervention and (2) discovery of ineligibility is not affected by the intervention or by prior knowledge of the researcher (Hollis and Campbell, 1999, Fergusson et al., 2002, Higgins et al., 2011a). Both conditions are fulfilled as the participant never started the intervention and her ineligibility was discovered by one of the blinded assessors without any participant data available to her/him. Therefore, and although mITT is not a proper ITT analysis and may bias the results, it can in the present trial be considered unbiased.

[108] Detection bias refers to systematic differences in outcome assessement (The Cochrane Collaboration, 2018).
[109] There is also a risk of selection bias inherent in research in general–see footnote 93.

Box 10 Risks of bias for effect results not considered in the Cochrane Collaboration's tool for assessing risk of bias (Appendix HH), by kind of bias and estimated direction, with estimated magnitude and origin of bias

Magnitude estimated as low, medium, or high. Origin of bias classified as: bias from trial design in general, bias from feasibility design in general (feasibility design), bias specific to this RCT which can or cannot be avoided in a full RCT (full RCT[–modifiable]), bias from this specific feasibility trial's design (specific trial design), bias from this specific feasibility trial's execution (trial execution), bias from design or execution error in this specific feasibility trial (error).

Selection bias

Unclear direction:

- Findings of small studies in a single centre are not representative (Guyatt et al., 2008). (medium, feasibility design) (H[a], unclear)
- The small feasibility sample size leads to a high likelihood of (unknown) imbalances in baseline participant characteristics and random variation (Arnold et al., 2009) with unknown residual confounding baseline imbalance and direction of bias (Jager et al., 2008). This leads to unreliable effect estimates (Arnold et al., 2009). (high, feasibility design) (H, unclear)
- Recommending therapy for urinary incontinence symptoms instead of accepting women into the trial in three cases might have introduced selection bias. (low, trial execution) (H, unclear)
- Because of the exclusion of one participant after randomisation, the sample considered for the clinical analysis slightly differs from the feasibility sample discussed in the process section. However, it still is not representative for Austrian women nor for the target population of childbearing women. (medium, trial execution) (H, unclear)

Towards overestimation of effect:

- Homogeneity of participants (Moore et al., 2011) (medium, trial execution) (H, unclear)

Performance bias

Unclear direction:

- The real-world use of the interventions cannot be compared as participating in a trial by itself may have a beneficial effect on the outcome through enhancing the level of attention, support and supervision (Hawthorne bias) (Braunholtz et al., 2001, Torgerson, 2014, Indrayan, no date). (high, trial design in general)
- Being a trial participant influences adherence, and follow-up in the trial was more intensive than in routine practice, with measures to improve adherence. (medium, specific trial design)
- Delay bias for delay between randomisation and intervention start (Torgerson, 2014). (low, trial execution)

Towards overestimation of effect:

- Adherence bias: Adherence was slightly higher in the experimental group. (low, trial execution)

Towards underestimation of effect:

- Instruction bias (Indrayan, no date) by the fact that experimental group participants may perform pelvic floor muscle exercises if they wish was not written in the participant information sheet. (low, specific trial design error)
- Not enhancing ball weight in the course of the trial did not lead to the exploration of the full potential of the experimental intervention and thus may have contributed to more gain in the comparison group. (high, specific trial design) (H, underestimation)
- Occasional inadvertent reminders of correct performance of pelvic floor muscle training during the adverse event phone calls might have intensified the comparison intervention. (low, trial execution error)
- 30 minutes of ball use including sitting times. (low, trial execution) (H, underestimation)
- Minimum standard care in Austria is the recommendation to perform pelvic floor muscle exercises. Although the intention in this feasibility trial was for the comparison group to receive standard care, participants in reality received enhanced standard care: they got extra instruction about pelvic floor exercises, evidence-based training instead of the gentle exercises usually recommended in Austria, had a professional digital vaginal check for correct pelvic floor muscle contraction, they documented their adherence and got adverse events calls which probably served as reminders. (high, trial execution)
- Contamination bias (Sackett, 1979): Contamination as “exposure to like treatments received outside the trial” makes it more difficult to find a treatment difference if one exists (Meinert, 2012, p. 93); none of the participants in the comparison group used a vaginal ball, but contamination by pelvic floor muscle exercises additionally to ball use was the case with about a third of the participants, half of them regularly (from three times/week to daily). However, the nature of the regular exercises probably comprised gentle exercises and not evidence-based training. (low, full RCT) (H, underestimation)
- Exercise regimens named for pelvic floor strengthening other than the scientifically recommended pelvic floor muscle training[110] used by participants: Four participants (two in each group) attended Pilates classes during the trial, 12 performed Yoga (in class or at home), and four attended Kanga classes). In a systematic review, Bø and Herbert (2013) concluded that there is not yet strong scientific evidence that exercise regimens other than pelvic floor muscle training can reduce stress urinary incontinence in women. Baessler and Bell (2008) argue that Pilates is likely to lack specificity and overload and therefore is not necessarily pelvic floor muscle training, but Culligan et al. (2010) show that it might

[110] Apart from the pelvic floor muscle exercises made widely known by Kegel (1948), other exercising techniques are used for pelvic floor rehabilitation purposes. These include e.g. Yoga, Tai Chi, Pilates, the Paula method, breathing exercises, abdominal muscle training, posture correction, and general fitness training (Bø and Herbert, 2013), but also Cantienica® (CANTIENICA AG, 2018), Antara® (Antara Training, 2018, Antara Training Österreich, 2018), Kanga (*Kanga Training*, no date), Kieser (Kieser Training AG, no date), and others (Franke, 2010). Except for Pilates which was shown to improve pelvic floor muscle strength in an exercise programme for women with little or no pelvic floor dysfunction (Culligan et al., 2010), no scientific evidence is available on these techniques (Bø and Herbert, 2013, own literature search).

nevertheless be helpful to strengthen the pelvic floor. As Baessler and Bell (2008, p. 209) write, ”[t]he integration of the pelvic floor into Yoga and Pilates may range from simple coactivation to voluntary and active incorporation of the” pelvic floor. Concluding from participants' accounts, the integration rather seemed the coactivation form. These techniques might be influencing variables when participants used them in this trial. Assuming Pilates to be effective, two Pilates participants in each group is biasing results towards more gain in the comparison group. (low, full RCT)

- Although women attending or planning to attend a postnatal exercise class at trial start were not included in the trial, one participant in each group attended such a class during trial participation. The attendance was minor: one started four weeks before the end interview (usual frequency once a week), one went three times during the trial. It can be questioned whether attending a postnatal exercise class once a week (often with the baby) without regular training at home can effectively contribute to pelvic floor muscle strengthening. The more influential variable might be women's regular pelvic floor muscle training at home, irrespective of participants attending a class, and it might be more important to consider this as postrandomisation mediator. A variable collected at the final interview that might influence the effect is performance of pelvic floor muscle exercises during the intervention period in the experimental group. (low, trial execution)

Detection bias

Unclear direction:

- A number of variables was self-reported. In this, errors could have occurred (recall bias) or socially desired answers may have been given (social desirability bias). (low, full RCT) (H, unclear)
- Some potential prerandomisation confounding variables for pelvic floor muscle rehabilitation such as further birth parameters (labour induction or enforcement, birth duration, epidural anaesthesia), levator ani avulsion or chronic pelvic floor stress (lifting, constipation, asthma) (Abdool et al., 2009a, Milsom et al., 2017, Salvatore et al., 2017) were not considered in this trial. This leads to unknown residual confounding baseline imbalance and direction of bias. (medium, full RCT–modifiable)
- Repeat testing bias (Indrayan, no date): In a pretest-posttest situation, participants might remember previous questions and may remove previous errors in the post test, thus do better without intervention effect (when asked, only few participants could remember their preintervention PRO values; one explicitly corrected one symptom to worse). Biological measurements have a tendency towards the mean. Also, the observer may acquire expertise to elicit correct response in the course of the study. (low, full RCT)
- During perineometry, the babies sometimes had been lying on or near the mother's chest. (medium, trial execution error)
- A few participants took pictures of their perineometry results and their knowledge of these results may have influenced their subjective pelvic floor ratings. (low, trial execution error)
- The answers' classification from ordinal to dichotomous could have introduced bias as different ways of categorising can give different results (Indrayan, no date) (low, specific trial design)

- ITT analysis biases towards no difference (Higgins et al., 2011b). (high, full RCT–modifiable)
- No sensitivity analysis was performed for variable outliers. (low, feasibility design)
- Data missing at random do not bias results, whereas data not missing at random risk to bias an available case analysis (Higgins et al., 2011b). The few missing data in this trial were mainly missing at random, only four values were missing not at random (see Appendix BB), and they were missing from different groups. (low, trial execution) (H underestimation)
- No sensitivity analysis was performed for missing data. (low, feasibility design)
- Some final measurements were delayed post intervention (low, trial execution)

Towards underestimation of effect:

- The ICIQ-UI SF sum score and thus the baseline urinary incontinence rate were higher in the comparison group; as the interventions might have more effect in weak muscles (Herbison and Dean, 2013), this might have biased effect results towards more gain in the comparison group. (medium, trial execution) (H, direction unclear)
- During lactation, oestrogen levels are low (Bonnar et al., 1975, Baird et al., 1979, Glerean et al., 2010), rising only with resumption of ovarian cyclicity (McNeilly, 1979). As oestrogen is needed by the lower urinary tract to function properly (Robinson and Cardozo, 2008) and also influences the striated muscle of the pelvic floor (Miodrag et al., 1988), less improvement may be seen with pelvic floor muscle rehabilitation during lactation than after its cessation and return of the menstrual cycle respectively (lactation/no ovarian cyclicity thus being a potential time-varying moderating factor). Of the 54 participants in the clinical analysis, one participant in the comparison group was not breastfeeding at trial start; additionally, two in the experimental and four in the comparison group (one of unknown status in each group) were not breastfeeding any more at the final interview, so that non-breastfeeding status at trial end was 2 (5.7%, one unknown) in the experimental vs 5 (27.8%, one unknown) in the comparison group. Five of the participants who had weaned during the trial had done so at least eight weeks before the final interview, and one two weeks before the final interview. (low, full RCT) (H, overestimation)
- Participants in the comparison group might be more able to perform the pelvic floor squeezes for the perineometry measurement at trial end as they are potentially more experienced in doing such contractions. (medium, full RCT)
- They might also be more able to perform the Knack and thus score lower on the ICIQ-UI SF. (medium, full RCT)

[a]H indicates that issue might also risk bias in harms results, with direction stated as this may differ from direction in effect results; topic is elaborated in the harms discussion section.

A promising effect size?

Moore et al. (2011) suggest that the decision to move forward to a full study shall not be based on the effect size from a pilot study. They point out that a homogenous sample within a single study site contributes to an overestimation of treatment effects, and that using a small estimate from a pilot study, such as that from the clinical outcomes in this feasibility trial, would result in a high likelihood of not investigating truly efficacious interventions. Arnold et al. (2009) stress the high likelihood of baseline imbalances due to small pilot sample size, which would lead to unreliable estimates of treatment effects.

Nevertheless, as stated in Chapter 4, one feasibility trial objective is to find out whether the calculated preliminary effect is "promising" and "worthwile to continue investigating". However, no guidance could be identified on the meaning of this statement. If a full trial's aim is to detect superiority of one intervention over another, some preliminary effect in the desired direction might be seen as promising; if finding no or a contrary effect is a satisfying result, too, finding a corresponding result in a preliminary trial might be considered promising.[111] As in this case any effect size would be appreciated in a future full (superiority) RCT, no effect or any direction of effect is considered "promising" in this feasibility trial.

Anticipated future RCT discussion

As this is a feasibility trial, the effect sizes are not discussed as they would be in a full RCT but with a focus on trial feasibility. However, attention needs to be drawn to a few issues which would be discussed in a full trial.

- Concentric/isometric exercises might not be the best form of pelvic floor muscle training.
- If, as Herbison and Dean (2013, p. 3) suggest, the cones' "effectiveness is likely to vary depending on […] initial pelvic floor muscle strength", with those having low strength having most to gain, balls might not be as effective for prevention than for treatment in incontinent women where weak muscles can reach a higher increase.
- It must be differentiated between level of adherence measured versus level of adherence needed for an effect. The most effective level of adherence is not known for either

[111] Excurs: To find no difference is the task of equivalence trials which try to find out if one intervention is not substantially worse or better than another one (Lesaffre, 2008). Similarly, a noninferiority trial asks whether the intervention under scrutiny is not substantially worse than a comparison intervention (Petrie and Sabin, 2009). Although noninferiority testing could be of interest for the present trial, it is not feasible. The reason is that in a noninferiority trial, the experimental intervention needs a comparison standard treatment which is based on sound evidence (Pocock, 2003); however, there is no research available with continent women starting preventive pelvic floor muscle training post partum. Furthermore, the three mixed prevention and treatment studies that looked at the outcome pelvic floor muscle strength measured by vaginal manometry (as by Woodley et al., 2017) found no difference between intensive training and usual care (Meyer et al., 2001, Hilde et al., 2013a, and Kou et al., 2013, the latter cited in Woodley et al. (2017)).

intervention. The adherence level was set for pelvic floor muscle training according to contemporary recommendation by the UK Association of Chartered Physiotherapists in Women's Health (2013b, 2013a) for routine postpartum pelvic floor muscle training; for ball use according to previous (not postpartum) cone trial protocols amended for pragmatic and logistical reasons and corresponding to popular use. According to Pelvic, Obstetric and Gynaecological Physiotherapy (POGP) (2017a), the British recommendations are based on the NICE guideline on female urinary incontinence (National Collaborating Centre for Women's and Children's Health, 2013). These recommend to perform a minimum of eight contractions three times per day without specifying their duration. The dose of the most recent evidence-based training recommendation for therapy is daily 1-3 sets of 8-12 contractions which are held for 6-8 seconds (Bø, 2015b, Bø and Aschehoug, 2015, Mørkved and Bø, 2015). However, there is no clear scientific evidence on the optimal training regimen, and the optimal training dosage for effective pelvic floor muscle training is not known (National Collaborating Centre for Women's and Children's Health, 2013, Mørkved and Bø, 2015). According to exercise science, the appropriate dose might already be reached with as little as performing exercise blocks on two to three (American College of Sports Medicine, 2009) or three to five (Haskell, 2012) days a week. Future research is needed to find the optimal training dose to treat and prevent urinary incontinence during childbearing (Mørkved and Bø, 2015).

- Measuring a good voluntary and maximum contraction in a supine position may not represent the actual function required of the pelvic floor (Bø, 1995a, Slieker-ten Hove et al., 2009) which is to work reflexively in the upright position.
- Other physiological aspects than pelvic floor muscle strength might be influenced by pelvic floor muscle training (Dietz et al., 2009). Likewise, vaginal pressure as determined by perineometry is only one facet of pelvic floor function and only an indirect measure of pelvic floor muscle performance (Peschers et al., 2001, Dumoulin, 2004, Bø and Sherburn, 2005, Bø, 2015c, Deegan et al., 2018). Other techniques for measuring pelvic floor muscle performance (Bø and Sherburn, 2005, Deegan et al., 2018) might bring about a different result.

Harms results

Screening for potential harms of the interventions was comprehensive by active surveillance and self-reporting of participants, resulting in 28 participants reporting a total of 38 adverse events (not considering recurrence). The reported adverse events and their frequencies correspond to the expected adverse event results named in the research protocol, with the exception of the frequency of vulvovaginal symptoms which occurred more often than the expected “very rare” vaginal irritations/infections. If the PP calculation of adverse events was

an mITT calculation (as advocated by some people (Higgins et al., 2011b)), the rates for muscle soreness and local discomfort/pain in the comparison group would be slightly lower.

Apart from one vaginal candidiasis/cystitis, vulvovaginal symptoms were the most severe adverse events and needed treatment. The four participants self-diagnosed their vulvovaginal symptoms as candidiases without consulting a medical doctor. Although women frequently do so, this allows for wrong diagnosis as the symptoms are not specific to the disease (Anderson et al., 2004, Johnson et al., 2010)[112]. The self-diagnosis of candidiasis also runs counter to the fact that the prevalence of candidiasis is lower when, as is the case in the postpartum time (Bonnar et al., 1975, Baird et al., 1979, Glerean et al., 2010, Jackson, 2011), sex hormones are reduced (Mylonas and Friese, 2010, Gätje et al., 2011, Gonçalves et al., 2016). However, the four participants self-treated their symptoms seemingly successfully (this might confirm candidiasis), although they recurred in one. The participant with the medically diagnosed and treated candidiasis/cystitis was 21 years old and can be assumed to have lacked experience with her body, compared to the other participants concerned who were between 29 and 38 years and with their longer life experience may have been able to treat their vulvovaginal symptoms themselves satisfactorily.

Vulvovaginal symptoms might possibly be attributed to mechanical irritation and resulting vaginitis by a foreign object as purported in gynaecology teaching texts (Gätje et al., 2011, Römer et al., 2012), although all four participants for whom the information is available were neither particularly adherent, nor did they use the ball substantially longer than 30 minutes. Potential postpartum vaginal atrophy through lack of oestrogen (Wisniewski and Wilkinson, 1991) may also play a role, particularly in lactating women (Bonnar et al., 1975, Baird et al., 1979, Glerean et al., 2010), which two of the four participants for whom the information is available were at trial end. According to manufacturer information (Intimina, no date), participants were instructed to clean the ball with the antibacterial soap provided. They however were not enquired about their adherence to this protocol; also, as mentioned by 12 participants (see section 6.6.1), the structure of the ball (indentations, surrounding plastic ring, textile retraction cord) may have hindered thorough cleaning.113 Other causes for the symptoms might be irritation by the lubricant or antibacterial soap used. The vulvovaginal symptoms rate compares well with Porta-Roda et al. (2015), who reported four of 37 participants in the vaginal spheres group with hypersensitivity, irritation, itching, and local

[112] Medically trained participants seem to be more able to recognise candidiasis correctly (Johnson et al., 2010)–this concerned one participant in the present trial (nurse).
[113] Intimina (2018f, https://www.intimina.com/en/laselle_weighted_exercisers) informs that the ball itself is “made entirely from body-safe materials, including phthalate-free silicone and ABS” (Acrylonitrile butadiene styrene (Encyclopædia Britannica, 2017)) and has a nonporous silicone skin, that the outer ring is of durable ABS plastic and that the retraction cord is bacteria-resistant.

discomfort. Bø et al. (1999) named two cases of vaginitis in 29 cone group participants; both trials studied nonpostpartum samples.

Muscle soreness, experienced by eight participants (15.1%, 95% CI [5.5, 24.7]), is a recognised training-related adverse reaction (Ratamess, 2012), and the NICE guideline describes occasional pain and discomfort as adverse effects of pelvic floor muscle training (National Collaborating Centre for Women's and Children's Health, 2013). The recorded muscle soreness not only in the comparison but also in the experimental group (for cones also reported by Fischer et al. (1996)) might indicate that ball use triggered muscle activity. Whereas participants in the experimental group only experienced pelvic floor muscle soreness, participants in the comparison group also experienced abdominal muscle soreness, which might result from concurrent pelvic floor and abdominal muscle contraction (Sapsford et al., 2001).

A third of the participants in the experimental group reported local pain or discomfort, particularly at ball insertion and removal, even leading to injury in one case; five and one participants respectively found the outer ring and the retraction cord of the ball uncomfortable. Pain or discomfort was also registered in other device studies. Porta-Roda et al. (2015) reported local discomfort among the adverse events experienced by four (of 37) ball group participants, Cammu and Van Nylen (1998) reported five participants with unpleasant feeling in their cone group with 30 participants. Similarly, Fischer et al. (1996), as only postpartum study, described discomfort by the inserted cone and the retraction cord for five participants, and Herbison and Dean (2013) mentioned discomfort as reason for dropping out of treatment. Of Glavind's (2001) 10 participants, two found it unpleasant to put the balls into the vagina and did not complete the study.

Pelvic floor muscle exercises had fewer uncomfortable adverse events reported. However, items named by five participants (of 17) was more often than in the comparable literature. In their systematic review on peripartum pelvic floor muscle training, Woodley et al. (2017) identified very rare pelvic floor pain. Lagro-Janssen et al. (1992) and Lagro-Janssen et al. (1994) (same study) described pain (once) and an uncomfortable feeling (three times) in approximately 88 participants. Being kept from falling asleep in the evening when performing the exercises before going to sleep, as named by one participant in this trial, was not identified in the consulted literature.

Vaginal discharge in the experimental group was a rare occurrence. Except by Kondo et al. (1995), who reported a rate of increased vaginal discharge of 2-10% (unclear) in their cone study, it is not mentioned in the literature. Haemorrhoids were reported to have worsened by

one participant, in contrast to one participant who reported improvement; haemorrhoidal changes discussed as (beneficial or adverse) cone effect could not be found in the literature.

Bø et al. (1999) reported one participant with bleeding out of 29 in their cone group. Of the four participants with nonphysiological vaginal bleeding in this trial, one attributed the bleeding to vaginal ball use. One participant had her vaginal bleeding clarified as not stemming from ball use by a medical examination. Other than that, no medical examination for the cause of bleeding was performed. Although this might explain the cause(s) of bleeding, it would be cumbersome to arrange a medical examination for every vaginal bleeding occurring in a full RCT; first because this would mean repeated examinations for participants who are back to their menstrual cycle, and second because this would put the burden of gynaecologist appointments (time and examination) on the participants. As this was only rare and slight bleeding of minor importance, participant report should also be relied upon in a full RCT. However, recurrence of bleeding should be enquired more carefully.

The causality relationship between an intervention and an adverse event, which converts the adverse event into an adverse reaction, needs to be concluded with care (European Medicines Agency, 1995). From a statistical point of view, the high rate of adverse events might be caused by chance, and a larger sample size in a future RCT could accumulate more scientific evidence to estimate if this was indeed the case. Many of the risks of bias collected for effect results in Appendix HH and Box 10 are also applicable to harms (in Box 10 indicated by H). However, the frequency of vulvovaginal symptoms confirmed the risk for vaginal irritation/infection reported in cone and ball studies (Bø et al., 1999, Glavind, 2001, Porta-Roda et al., 2015), and irrespective of a medical diagnosis, the frequency of symptoms should attract attention. It is remarkable that, except for one withdrawal, all participants continued participation and ball use in spite of the adverse events experienced.

A limitation to the adverse events screening performed is that the screening system was specifically designed for this trial and had not been piloted before (in line with the feasibility trial design). From a methodological point of view and as described in the respective trial processes section (6.5.3), the probing process for adverse events therefore was suboptimal and might bias the results towards missing harms (except for vulvovaginal symptoms and infection); also, event recurrence was not collected explicitly but elaborated retrospectively from the documented information. As with effect results, generalisation of harms findings to the target population is difficult.

10.2.4 Participants' experiences with and opinion on the trial

Overall, the trial experience for the participants seems to have been positive. Many called participation uncomplicated, interesting or exciting, and they found it enhanced their

motivation to do something for the pelvic floor. As benefits, they named the learning experience, pelvic floor improvement, the emotional benefit of contributing to research and thereby helping others (reflecting their original altruistic motivation for trial participation), the technical pelvic floor measurement per se and being quick and easy, and this all at no costs. The good experience and the benefits named correspond to the findings of Brubaker et al.'s (2013) focus group study on participation experience in pelvic floor studies. Reflecting on these authors' results, an implicit benefit in this trial also seemed to be participants' quality relationships with research staff, as concluded from their thanking comments and feedback that they had been well cared for. As pointed out to be of interest to participants by Partridge and Winer (2002), participants in this trial also asked about other participants' experiences.

It became clear however that trial participation also could be a burden. This comprised the sense of an additional obligation, the (sensed) requirement not to attend a postnatal exercise class, and time or adherence issues, but went as far as feeling guilty and being afraid of further ball use after vulvovaginal symptoms (the last two points were not mentioned as trial experience but when asked about the intervention). Developing hitherto unknown worries about the pelvic floor and vulvovaginal symptoms as potential adverse effects of ball use were also named as disadvantages of participation. The issues identified partly correspond to the answers in Brubaker et al. (2013) who named the required commitment and complications as disadvantages. However, negative experiences and disadvantages were named by 13 participants only, and as for the participants in Brubaker et al.'s (2013, p. 77) study, it seemed that negative experiences were "minor disappointments or inconveniences but not 'deal breakers'".

Although the experiences are encouraging towards a full RCT, future potential participants should be alerted to the potential burden of trial participation, citing past participant experiences. Of the suggestions to enhance the trial, deleting weekday names from the adherence sheets' heading line had already been resolved during the course of the trial. A measurement comparison value from before pregnancy, wearing the ball for longer than 30 minutes, and ticking adherence instead of noting minutes on the form do not seem realisable. The other suggestions were discussed in the respective sections.

Summary

Despite some difficulties, the recruitment rate at the researcher level was high enough. Women's high motivation to participate and to perform their allocated intervention might stem from their wish to do something for the pelvic floor. The feasible selection criteria need minor modifications, e.g. the exclusion of women with urinary incontinence symptoms. The highly educated and socially well-situated opportunity sample is not representative for childbearing

women, and in a full RCT, strategies need to be applied to enhance sample representativeness. The comparably low attrition rate must be seen with caution.

Organising perineometry appointments and fulfilling Feasibility criterion 5 might be facilitated by more convenient appointment times and childcare service during the perineometry measurements. Perineometry data collection and harms screening need some improvements, and collection of adherence data might be enhanced by electronic data collection forms. The administration of the online survey was shown to be feasible, but its content needs to be adapted to the full trial.

Better informing participants on a timely intervention start would serve Feasibility criterion 2. Despite some difficult aspects, both interventions were feasible. However, with a maximum adherence rate of 62.9%, Criterion 4 was not fulfilled. In a full trial, measures to enhance adherence should be applied in both groups; this does not only concern applying a high enough dose of the intervention but also not surpassing the recommended dose (usage time of ball). Ball handling after vulvovaginal symptoms needs to be reviewed for a full trial.

All management experiences from this feasibility trial can be used to optimise a future trial. Because of funding issues however, some findings might be specific to the PhD circumstances of the project and not be fully applicable to other research settings. Resources needed for this feasibility trial were compiled and potential for higher and lower costs for a full RCT were identified.

The calculated change score differences between the groups are small and have broad CIs, pointing to uncertain results or potentially even showing a contrary effect; also, the direction of effect results is partly contradictory. The results regarding adverse events point to potential risks, whereby the broad CIs show uncertainty as well. Both effect and harms results need to be regarded with caution as the feasibility trial features and other study characteristics might lead to bias. This could result in an over- or underestimation of effect and/or harms, and in accordance with a feasibility research design, the generalisability of these clinical results therefore is low.

The participants' opinion on the trial encourages a full RCT and gives input for future trial planning. After this discussion of the specific feasibility results, the next chapter reflects on implications for future research.

11 REFLECTIONS ON AND IMPLICATIONS FOR FUTURE RESEARCH

Similar to methodological triangulation as combination of “different research strategies in the study of the same empirical” unit (Denzin, 2009, p. 308), the integration of the results from the different result areas informs the implementation and conduct of future research and in particular a full RCT. This chapter starts with a reflection on whether a future full RCT should be performed at all, and thereafter, (potential) implications for future research are considered.

11.1 A future RCT?

Although knowledge on the effect of hold and vibration use of vibrating vaginal balls to improve pelvic floor muscle performance in the postpartum period was gained in this feasibility trial, clinical equipoise and thus the rationale for a full RCT are still present. The effect size CIs are wide and include no effect or an effect in the other direction, and a full RCT could provide greater certainty to the effect estimates (Eldridge and Kerry, 2012). Furthermore, a full trial would perform significance testing for (at least) the primary outcome to determine whether the data are likely to have been found by chance ("unexpectedness test", Clark-Carter, 2003, Mark et al., 2016), and with the calculated *p* value perform null hypothesis testing (Mark et al., 2016)[114]. Process and management results confirm that, with modifications, a future RCT might be feasible.

For the different reasons specified on pages 16/17, the use of vibrating vaginal balls might be seen by women as advantageous to pelvic floor muscle training, or their use could influence outcomes other than pelvic floor muscle strength (as mentioned in Footnote 36). However, in view of the adverse events encountered, it first needs to be clarified whether it is ethically justifiable to progress to a full trial. The fact that women in the experimental group had more problems with discomfort and vulvovaginal symptoms/infections suggests that vaginal balls may not be appropriate for women in terms of an intervention for a future larger RCT, the more when seen in combination with the fact that the comparison group showed better preliminary results for most of the effect outcomes. To come to a conclusion in this matter, the question about acceptable harms should be discussed by the research team and women in PPI work[115] or be researched in an appropriate feasibility study (e.g. survey or interviews

[114] However, controversy exists about the value of null-hypothesis significance testing (Grissom and Kim, 2012, Mark et al., 2016).

[115] PPI could, as in this feasibility trial, happen at the level of consultation with women being consulted on key aspects in trial planning and preparation. It could also, at a higher level, happen as collaboration–an ongoing partnership between researchers and PPI participants as members of the trial management group througout the research process (INVOLVE, 2012). Bagley et al. (2016) provide a toolkit to support trial teams in undertaking PPI, an online resource hub is provided by Imperial College London (2018).

of potential participants). Another important question to be clarified within PPI is whether (a statistically significant result of) the primary outcome pelvic floor muscle strength, unlike symptoms not of direct clinical significance, is of practical relevance to women.[116]

For the PPI meeting(s), a draft research plan for a future RCT must be presented as a basis for discussion. This plan already needs to have necessary protocol modifications incorporated. If the PPI participants together with the researchers come to the conclusion that a full RCT is not justified, the project will be aborted. If PPI participants agree that a full RCT is justified and of interest to women, further PPI input will be needed to refine and agree upon the presented research protocol draft; this shall ensure a viable trial protocol which is adapted to the future research setting and to which women agree. Using input from all trial results, relevant issues to consider in planning a future RCT are elaborated in the remainder of the chapter.

11.2 Implications for a potential future RCT

Points to consider for a full trial protocol are given under the next headings. Implications are named for methods and design, sample size calculation, processes, management, and resources. As a trial needs to be tailored to its context, the suggestions are generic and open to adaptations. Assessment and evaluation of the presented solutions, as suggested by Bugge et al.'s (2013) framework, need to be performed by the future trial team.

11.2.1 Methods and design issues

Methods and design issues for a full RCT concern the question of a pragmatic versus an explanatory research interest, the participants, the choice of interventions and primary outcome measure, the analysis, and the trial design.

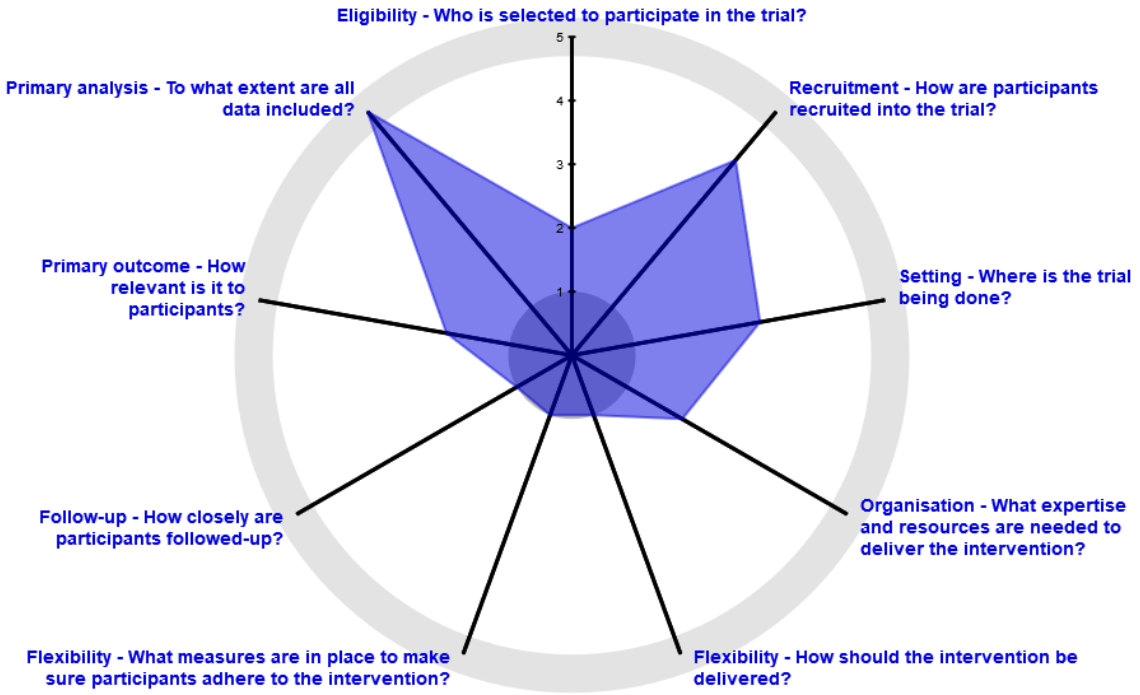
Pragmatic or explanatory research interest

So far in this thesis, the term "effectiveness" has been used without explicit definition to denote an intervention effect. However, a difference was pointed out between effectiveness (or pragmatic) and efficacy (or explanatory) trials by Schwartz and Lellouch (1967). Both kinds of trials try to determine the effect of an experimental treatment, but the former is performed under everyday conditions and investigates whether the intervention does work when offered as therapy, whereas the latter is performed under ideal experimental conditions and tests causal hypotheses (Schwartz and Lellouch, 1967, Eldridge, 2010). However, as Meinert (2012, p. 430) states, "there is no clear line of demarcation between the two".

[116] Kirk (1996) and Thompson (2002) differentiate between statistical, practical and clinical significance of results.

Thorpe et al. (2009) contributed to the pragmatic-explanatory differentiation in that they suggested not to look at pragmatic and explanatory as binary ends of a continuum for the whole trial, but to consider a differentiation within nine domains. For this purpose, they introduced the Pragmatic Explanatory Continuum Indicator Summary (PRECIS) tool, in the meantime developed into the PRECIS-2 (Loudon et al., 2015). This tool was applied to this feasibility trial after its completion, and Figure 28 and Table 25 show the evaluation of the domain characteristics. This indicates a primarily explanatory trial design.

Figure 28 PRECIS-2 wheel for present feasibility trial



According to Loudon et al. (2015). The PRECIS-2 tool predefines the areas of differentiation between a pragmatic and explanatory design; the centrum is the explanatory pole, the outer ends of the lines are the pragmatic poles. Depending on where on the continuum a particular trial is assumed to be located within a domain, a mark is put on the line.

Table 25 PRECIS-2 scoring (Loudon et al., 2015)

PRECIS-2 domain	Score[a]	Rationale
1 – Eligibility criteria	2	• Eligibility criteria possibly comprised less than half of the women who could use the device in practice (by e.g. only including women after vaginal term birth, over the age of 18, with sufficient knowledge of German and excluding those currently enrolled in pelvic floor muscle training with professional, 3rd or 4th degree perineal tear at most recent birth). • Women after caesarian section and preterm birth were excluded as not expected to be highly responsive to the intervention. • Women who planned to move away during the study period (as women whose follow-up might have posed difficulties) were not included.
2 – Recruitment	4	• Most participants recruited with the help of their healthcare professionals from usual care appointments at different sites. • Additional recruitment by flyers and websites (used to speed up recruitment and to try out recruitment routes). • Travel costs to measurement site (partly) refunded.
3 – Setting	4	Single city
4 – Organisation	2	• No healthcare staff and resources necessary for interventions as self-administered by participants, but material provided for free which in usual care participants would need to purchase themselves. • Costs for collection of data as these would not be collected as part of usual care.
5 – Flexibility (delivery)	1	• Highly specified, protocol driven interventions with specific directions for how to administer them (instruction, dose). • Measures in place to monitor adherence with the protocol and to address poor adherence. • Restriction on cointerventions: women currently enrolled in pelvic floor muscle training with

PRECIS-2 domain	Score[a]	Rationale
		professional were not included into the trial.
6 – Flexibility (adherence)	1	• Being trial participant may have affected adherence (starting/engaging with interventions). • To monitor adherence, participants completed an adherence sheet which would not be part of usual care. • They also received adverse events calls which may have enhanced adherence.
7 – Follow-up	1	• Intervention related data would not be collected outside the trial. • Trial visits were not part of usual care and involved additional, different staff.
8 – Primary outcome	2	• Pelvic floor muscle strength, a parameter thought to influence pelvic floor symptoms. Participants and healthcare professionals are rather, but not exclusively, occupied with symptoms. • Use of an assessment method not normally used in usual care and requiring special training.
9 – Primary analysis	5	mITT using all available data

[a]1 = very explanatory, 2 = rather explanatory, 3 = equally pragmatic and explanatory, 4 = rather pragmatic, 5 = very pragmatic.

As this trial's results provide scientific evidence that vaginal balls may not be suitable for or chosen by all women post partum, the effectiveness of informing all women on the choice of a vibrating vaginal ball to strengthen their pelvic floor (the pragmatic question of offering the intervention) turns out to be rather irrelevant. Of more interest is the question whether the balls work for women who would like to use them. This means that the research interest behind this feasibility trial and the purpose of a full RCT are shifting towards the explanatory end, with the outcome of interest becoming intervention efficacy. This fits well with the evaluation results of the PRECIS-2 scheme, and it therefore is suggested to explicitly design the future RCT as an explanatory trial.

Participants

As discussed in section 0, it is recommended to only include urinary continent women (ICIQ-UI SF sum score of 0) in a full trial. This would enable the trial to answer the research question but would not keep those from therapy who, according to scientific evidence, are

recommended treatment (Dumoulin et al., 2017). Enrolling only continent women would also correspond to the explanatory design because of tightening the inclusion criteria (Loudon et al., 2015).

It is also recommended to include women attending or planning to attend a postpartum exercise class who were excluded from this trial. Attending such a class once a week is unlikely to influence pelvic floor muscle performance, as from an exercise science perspective exercise blocks must be performed on at least two to three days a week (American College of Sports Medicine, 2009).

Choice of interventions

By using only one ball of 28 g, the experimental group in this feasibility trial practised a simple version of the intervention under study. Thereby, standard care was not withdrawn: Participants were neither discouraged from nor encouraged to perform the pelvic floor muscle exercises they had been recommended by professionals but informed they could do them if they so wished. In Austrian standard postpartum care, women are, as was the case in this trial, usually provided with written information on gentle pelvic floor muscle exercises for the early postpartum time and not with evidence-based muscle training information. The comparison group in this feasibility trial received enhanced Austrian standard care as the participants got the British postpartum pelvic floor muscle training recommendations which are based on exercise science.

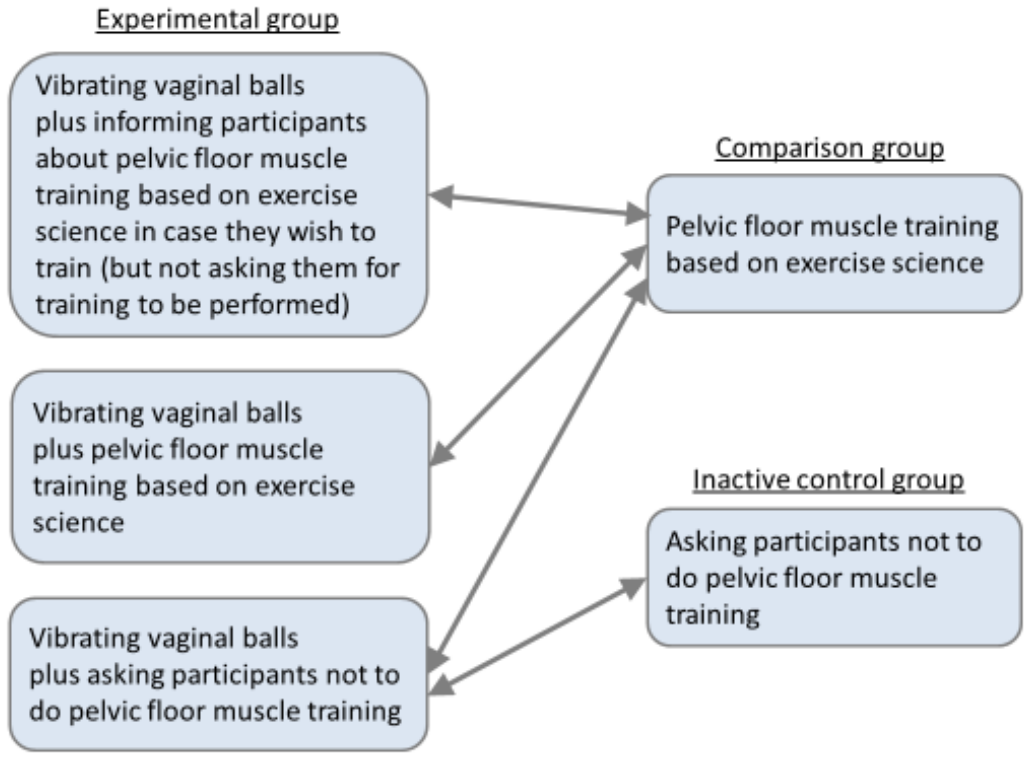
A future full RCT might profit from changes with respect to the interventions. First, the applied experimental intervention should be modified to use balls of increasing weight, starting with the heaviest weight each participant can retain. Second, the applied standard pelvic floor care needs to be reconsidered. Routine pelvic floor care post partum differs between countries and can e.g. be the recommendation to perform potentially ineffective gentle pelvic floor muscle exercises or supervised pelvic floor muscle training based on exercise science. As for research, interventions should follow the highest scientific standard (which is assumed to seep into local practice) rather than current local practice, standard care for the comparison group in this trial was enhanced to correspond to evicence-based pelvic floor muscle training principles; this should be kept in a full trial. However, in a full trial, evidence-based pelvic floor muscle training principles should be introduced to the experimental group as well to avoid performance bias by different standards of pelvic floor muscle training in the trial groups. As in this feasibility trial, intervention duration will be 12 weeks, still following the reasoning provided in section 5.3 (page 88).The respective recommendations must be adapted to future pelvic floor muscle training guidelines which might change with evolving scientific evidence.

Different intervention and comparison options for the trial groups in a full trial are shown in Figure 29. The experimental group may only be informed about evidence-based pelvic floor muscle training and the decision to train left to each participant, so that *if* participants train they do it according to the highest standard. Another option would be to use vibrating vaginal balls as an addendum to performing pelvic floor muscle training. This design of adding a new to the standard intervention in a combination therapy arm is suggested by Stanley (2007, p. 1165) for new drugs that look “promising but are not as potent as the current standard”. This however would impose a higher “intervention burden” on the participants since such a group would have to do both interventions, and it is therefore questionable if this would be feasible. Also, the research interest so far is to use the vibrating vaginal balls instead of pelvic floor muscle training and not as an addendum.

As intensive pelvic floor muscle training starting in pregnancy has been shown to decrease the prevalence of urinary incontinence at three months and up to six months post partum when compared with usual care (Woodley et al., 2017), it is expected that this intervention will seep into routine practice (see e.g. the APPEAL project (National Institute for Health Research CLAHRC South West Peninsula, 2018)). This means that potential future participants might be more likely to perform pelvic floor muscle training on their own, reinforcing the need to explore the mediating effect of regular exercising in the experimental group.

As preventive pelvic floor muscle training starting after birth to date has not been shown to be effective (as not researched), and mixed prevention and treatment have shown an uncertain effect (Woodley et al., 2017), it might, from an equipoise and ethical point of view, be possible to conduct a comparison of using vibrating vaginal balls plus asking participants to refrain from pelvic floor muscle training in an experimental group versus asking participants to refrain from pelvic floor muscle training in a control group. However, as women who have started pelvic floor muscle training in pregnancy can be assumed to continue their training post partum and to exercise more than women not having started pelvic floor muscle training during pregnancy, such a comparison would only be feasible with women who have not started pelvic floor muscle training in pregnancy, or would need adjustment for this factor. Alternatively, the number of trial groups could be increased to three (experimental, comparison, inactive control), again recruiting women who have not started pelvic floor muscle training in pregnancy or adjusting for this.

Figure 29 Options for interventions (boxes) and comparisons (arrows) in future RCT groups



Most appropriate primary outcome measure

An intention of this feasibility trial was to identify the most appropriate primary outcome measure for a full RCT. The primary end point of a study is "the outcome measure used to make the decision on the overall result of the study and serves as the basis to determine the number of patients needed for the study" (Stanley, 2007, p. 1167). Of the outcomes tested in this trial[117], the most appropriate primary outcome must answer the research question and can be selected in terms of assessment feasibility and reliability, statistical issues, and relevance for participants.

The research interest of a full RCT will be a gain in pelvic floor muscle strength, the statement made about ball use (FUN FACTORY, no date-d) and the argument of midwives' postpartum routine recommendations. As the future research question about vibrating vaginal balls (see section 11.2.2) therefore is on strengthening the pelvic floor (and not enhancing pelvic floor symptoms), measures of pelvic floor muscle performance–participants' self-reporting and perineometry–were chosen for this feasibility trial and are appropriate primary outcome measures in a future RCT.

[117] This does not preclude, in principle, choosing an outcome which was not tested in this trial (outcome options summarised by Deegan et al. (2018)).

Of the perineometry measures, resting pressure to date has not been given much attention in research, and the operationalisation of endurance measurement used in this trial might not be reliable enough. This leaves contraction strength pressure as the most used and researched perineometry measure. Additionally, more experience has already been gained with this measure than with participants' self-reporting of pelvic floor muscle strength. When comparing the two measures from a statistical viewpoint, maximum perineometric strength shows, in this feasibility trial, a (Pearson) correlation of before and after intervention values of r = .83 (95% CI [.67, .99], n = 53, corresponding to r = .84 in Myer et al. (2018)), compared to r = .45 (95% CI [.20, .71], n = 53) for participant reported strength. But, whereas perineometry is an intersubjective (see Footnote 75) measurement method, PROs represent women's experiences of their bodies. According to the International Consortium for Health Outcomes Measurement (2018b), it is desirable to choose outcomes that are important to participants (one participant in this study even suggested to put more weight on subjective pelvic floor perception). Thus, the decision on the best primary outcome is also one that can be taken by the future research team together with PPI participants.

Adherence and effect analysis

In case the comparison intervention stays identical, the adequate adherence threshold in a future RCT could remain 80% but the calculation should apply the wider adherence criteria developed in this feasibility trial analysis. However, as in other studies (e.g. Mørkved and Bø, 1997, Chiarelli and Cockburn, 2002, Stafne et al., 2012) and according to exercise science (as laid out on page 219) and future pelvic floor muscle training research on optimal training dose (Mørkved and Bø, 2015), training at least three times a week as comparison intervention might become considered as adherent. In this case, calculation of adherence might need to be reconsidered. A connected question is whether, when a smaller dose of pelvic floor muscle training will be recommended, the frequency of ball use also needs to be adapted to this lower dosage.

For effect analysis, a systematic statistical analysis plan as presented by Gamble et al. (2017) is strongly recommended for a full trial protocol. For the repeated measures design of the trial, the statistical analysis of the clinical outcome data should be performed by mixed modelling. This form of analysis can best take account of the three repeat perineometry measurements of pelvic floor muscle strength per participant each before and after the interventions (Petrie and Sabin, 2009). In contrast to a repeated measures ANOVA as a possible alternative, it is also more insensitive against missing values and able to analyse

repeated measures of response variables that do not follow a normal distribution (The Analysis Factor, 2018, The MathWorks, 2018)[118].

Sensitivity analyses need to be planned to test the robustness of the findings against the impact of baseline imbalance, missing data, variation of distributional assumptions, outliers, and different forms of analysis (Thabane et al., 2013). As adherence as mediating variable needs to be considered in the analysis of an intervention (Efron and Feldman, 1991, Goetghebeur and Shapiro, 1996, Cuzick et al., 1997), an explanatory adherence sensitivity analysis in the form of a PP analysis should be performed. Although such a PP analysis introduces attrition bias as the compared groups might not be similar any more (as not analysed by randomisation but influenced by participant characteristics), it better reflects effects of treatment when these are applied as planned in the research protocol ('Intention to treat analysis and per protocol analysis: complementary information', 2012). This PP analysis can be compared with the ITT analysis, and Meinert (2012, p. 262) suggests that ITT and PP analysis together "provide an interval estimate of effect". Another postrandomisation mediating variable which needs to be considered in a full trial analysis is pelvic floor muscle training performed in the experimental group. The return of ovarian cyclicity might be added to the model as a moderating variable. Missing values will be dealt with using the technique appropriate to the reason why they are missing; for continuous data not missing at random, multiple imputation is a recommended method (Higgins et al., 2011b, Kang, 2013).

Trial design

When progressing to a full trial, an internal pilot trial is suggested as the next step in the evaluation process of the researched intervention. Testing the full trial protocol is needed as new feasibility issues might arise with modifications of the feasibility trial protocol tested in this study, and in other research settings. Internal pilot trials (as opposed to external pilot and thus feasibility trials) are also suggested by Cooper et al. (2018) who found high variability when comparing randomisation and attrition rates between external pilot and and their associated full trials. If the internal pilot trial runs smoothly, the data collected could be carried forward to the full trial dataset (National Institute for Health Research, no date-a); Charlesworth et al. (2013) provide a systematic approach for decision-making on whether this should be done. A design option to consider is a preference trial.

[118] Alternatives to the frequentist framework used could be a Bayesian design and analysis strategy (Pezeshk, 2003) or hybrid frequentist-Bayesian approach (O'Hagan et al., 2005); a respective discussion is complex and outside the scope of this thesis.

11.2.2 Sample size calculation for a full RCT

It is unethical to unnecessarily expose persons to the risk of research, as would be the case with too many or too few planned participants (Lenth, 2001, Leon, 2008). An a priori sample size calculation therefore is mandatory for clinical trials (Schulz and Grimes, 2005). Although the main interest of an RCT needs to be on effect sizes, precision and clinical importance of effects (Cumming, 2013), the aim of a full trial is to detect a clinically significant effect with reasonable confidence, at the same time allowing for acceptable errors (Guyatt et al., 2008).

With hypothesis testing as the focus of a future trial[119], the required sample size calculation components are the hypothesis framework (superiority or noninferiority), the statistical significance test[120] going to be used, the desired study power and significance level, the assumed minimal clinically important difference (MCID, for superiority design), and the variability of the outcome values (Leon, 2008, Petrie and Sabin, 2009, Noordzij et al., 2010). The hypothesis framework and the MCID are considered in the following, before the sample size will be calculated.

Hypothesis framework

The future RCT's research question would be (subject to intervention change):

> Does using vibrating vaginal balls differ from performing pelvic floor muscle training in its effect on pelvic floor muscle performance from pre- to postintervention measurements in women after childbirth?

This research question is transformed into the two superiority design trial hypotheses shown in Table 26. The hypotheses of the future trial are two-sided hypotheses where the direction of the effect is not specified, allowing for either eventuality (Petrie and Sabin, 2009).

[119] Other possible trial foci are determining the precision of an estimate or width of a CI (Petrie and Sabin, 2009).
[120] With enhanced statistician input, sample size planning can be aligned with results reporting by basing it on estimated effect size CIs instead of significance tests (Bland, 2009).

Table 26 Trial hypotheses in future RCT (inspired by Lesaffre (2008), Petrie and Sabin (2009), Walker and Nowacki (2011))

Null hypothesis H_0 Tested statistically Rejected/not rejected[a]	States that there is no difference between interventions ($\Delta = 0$): • Pelvic floor muscle strength change from pre- to postintervention is the same in both groups.
Alternative hypothesis H_A Accepted if H_0 rejected	States that there is a difference between interventions ($\Delta \neq 0$): • Pelvic floor muscle strength from pre- to postintervention is statistically significantly different between the groups.

Note. Δ = difference between interventions results.
[a]It is important to note that, with hypothesis testing in a superiority framework, the result of no difference between the groups (the null hypothesis) can never be *accepted* as there might be a difference which simply has not been found (yet) (Barker et al., 2002, Walker and Nowacki, 2011).

The minimal (clinically) scientifically important difference

The future trial analysis method would be mixed modelling; as a mixed model sample size estimation requires enhanced statistical calculations, a *t*-test for two independent means as simplified option is used here for a first conservative estimation. Therefore, the minimal important difference that needs to be determined for the present sample size calculation is the difference in pelvic floor muscle strength (by perineometry or participant-rated) change scores between two postpartum intervention groups. Both potential primary outcomes in this research are clinical in the sense that they involve the direct testing of participants ('Definition of 'clinical'', 2018), but no professional judgement on a clinically reasonable difference can be made because there are no standard reference values that need to be attained. Instead, a *scientifically* important difference (MSID) must be identified (Lenth, 2001).

To determine an MSID, professional reasoning and data from earlier trials can be used (Halpern et al., 2002). Although Kraemer et al. (2006) caution to base the magnitude of the desired treatment effect on pilot result data because of pilot study effect sizes being biased due to small sample size, feasibility results are a “realistic estimate based on emerging empirical evidence […] about the likelihood of a particular effect size” (Guyatt et al., 2008, p. 4). In this sense, this work has provided a preliminary estimate of attainable change scores in the potential primary outcomes of a future full RCT. In the following, this feasibility trial’s results are, together with information from earlier trials, considered to estimate which effect size is reasonable and possible to expect in a full RCT.

Box 11 and Box 12 show the information and reflections supporting the decision on the MSID and the estimated MSID for each potential primary outcome of a full trial. The ideal earlier trial to be consulted for assumed MSIDs and effect sizes found would have a) used cones or pelvic floor muscle training for b) continent women c) within six months post partum, and d) measured this by perineometry with the Peritron or with the collected PROs. As no such trial exists, the closest comparable trials were consulted. To clarify the limitation of their contribution, their characteristics are provided.

Box 11 Perineometry strength pressure: information gathered to decide on the MSID, followed by MSID estimation

Assumed MSIDs in earlier RCTs[121]

Culligan et al. (2005):

- Participants: Primiparae, any type of delivery, mixed continent and incontinent, within 12 months post partum
- Interventions: active versus sham extracorporeal magnetic innervation to restore pelvic floor muscle strength
- Measured by perineometry (not by the Peritron) in cm H_2O, the assumed MSID is ≥ 30% for active versus sham (no intervention) group. Apart from the fact that the authors' decision was based on an ealier trial by Mørkved and Bø (1997), no more detail on this decision could be found out.

Effect sizes found in earlier RCTs

Hilde et al. (2013a):

- Participants: primiparae with singleton vaginal birth from 32 weeks of pregnancy onwards, mixed prevention and treatment of urinary incontinence
- Interventions: training class plus diary versus encouragement to perform training; all participants got a customary information leaflet, thorough initial instruction, and checking of correct contraction
- Measured by manometry (not by the Peritron) six weeks and six months post partum, strength change scores were 15.7 cm H_2O in the training group and 12.1 cm H_2O in the control group, the change score difference could therefore be calculated as 3.6 cm H_2O.

Effect size (change score difference) found in this feasibility trial: 4.6 cm H_2O

MSID reflections and estimation

- Bø et al. (1990a) measured a pelvic floor muscle strength increase of 100% in the first month of pelvic floor muscle training in urinary incontinent participants. According to the American College of Sports Medicine (2002), muscles have a strength gain of approximately 40% in untrained participants, of 20% in moderately trained, and of 16% in trained participants (over periods from four weeks to two years). Judging continent postpartum women as moderately trained, in contrast to assuming the incontinent participants in Bø et al. (1990a) as untrained, 50% as half the pelvic floor muscle strength increase is assumed for MSID calculation in this trial.
- The initial overall pelvic floor muscle strength mean in this feasibility trial was 23.0 cm H_2O; calculating an increase of 50% from this mean results in an absolute increase value of 11.5 cm H_2O; 30% (the relative MSID) of 11.5 cm H_2O is 3.5 cm H_2O. This aligns well with the effect sizes of 3.6 cm H_2O found by Hilde and the (biased) 4.6 cm H_2O found in this feasibility trial. The MSID between groups is thus estimated as between 3.5 and 4.6 cm H_2O.

[121] Although Ahlund et al. (2013) made a power calculation with the outcome pelvic floor muscle strength in a trial on pelvic floor muscle training post partum and measured this by the Peritron, their values are not considered as they are much higher than the values used here (their values, converted from cm Hg, are: MSID 95.2 cm H_2O, change scores 133.3 and 82.9 cm H_2O respectively).

Box 12 Participant reported pelvic floor muscle strength: information gathered to decide on the MSID, followed by MSID estimation

Assumed MSIDs in earlier RCTs
None
Effect sizes found in earlier RCTs
Other trials: none Effect size (change score difference) found in this feasibility trial: 5.1%
MSID reflections and estimation
• Aligning the calculation for this subjective measure with the calculation for the technical measurement in Box 11 leads to the following reflection: The initial overall participant reported pelvic floor muscle strength mean in this feasibility trial was 75.3%; calculating an increase of 50% from the mean results in an absolute increase of 37.7% and a final mean of 113%; 30% of 37.7% is 11.3% (20% alternatively is 7.5%). This is higher than the (biased) effect size of 5.1% found in this feasibility trial. The MSID between groups is estimated as between 5.1% and 11.3% cm H_2O.

Sample size needed

The ingredients for the sample size calculation for a fully powered RCT developing out of this feasibility trial are presented in Table 27. A two sided superiority hypothesis and a *t*-test for two independent means are used for a simplified first conservative estimation. Significance level and power are set according to convention as .05 and 80% or 90% respectively (Petrie and Sabin, 2009). The two potential MSIDs are taken from the calculations in Boxes 11 and 12 above, and the variability of outcome measures from this feasibilty trial's clinical results. A sensitivity analysis with respect to desired power and both potential primary outcomes (Matthews, 2000) is performed. Further, both a simple calculation and a calculation assuming a 15% attrition rate (higher than in this trial for different context) are considered. However, to make the final decisions for the sample size calculation for a future RCT, PPI work as mentioned in section 11.1. is needed to identify the most appropriate outcome measure from women's point of view.

The power analysis was performed via the programme G*Power (Buchner et al., 2013). The result from the sample size calculations for the two potential primary outcomes is shown in Table 28 and lies between 104 and 773. Limitations to this calculation are the nonrepresentativeness of the *SD*-variance originating from a nonrepresentative sample (Lenth, 2001), the difficulty to estimate an MSID, and the simplified preliminary calculation by the *t*-test. Having more than two trial groups or a planned subgroup analysis would increase the necessary sample size and need adjustment for multiple comparisons (Petrie and Sabin,

2009, Meinert, 2012, Wason et al., 2014). For a future funding application, an elaborated mixed model sample size calculation needs to be performed.

Table 27 Ingredients for sample size calculation for a fully powered RCT with two groups

Element	Characteristic in this trial	Choice based on
Alternative hypothesis H_A	Two sided superiority hypothesis	• Research question • Hulley et al. (2013)
Future analysis model	Mixed modelling; *t*-test for two independent means used here for simplified first conservative estimation	• Petrie and Sabin (2009) • Statistical advice
Desired power	80% or 90% (by convention)	Petrie and Sabin (2009)
Desired significance level (α)	.05 (by convention)	Petrie and Sabin (2009)
MSID	• For perineometry squeeze pressure (strength): 3.6 or 4.6 cm H_2O • For participant reported pelvic floor muscle strength: 5.1% or 11.3%	Professional reasoning as laid out in Box 11 and Box 12
Variability of outcome measure	• Perineometry squeeze pressure (strength) change scores (*SD*s): Experimental group: 7.7 (12.2) Comparison group: 3.1 (5.4) • Participant reported pelvic floor muscle strength change scores (*SD*s): Experimental group: 9.0 (15.9) Comparison group: 14.1 (24.0)	Feasibility trial data

Table 28 Preliminary total sample size estimations for future RCT, simple and corrected for attrition (+ 15%)

Primary outcome	Sample size needed for trial with power of			
	80%		90%	
	Simple	Corrected	Simple	Corrected
Perineometry squeeze pressure (strength) MSID 3.5 cm H_2O	232	267	308	354
Perineometry squeeze pressure (strength) MSID 4.6 cm H_2O (found in this trial)	134	154	180	207
Participant reported pelvic floor muscle strength MSID 5.1% (found in this trial)	504	580	672	773
Participant reported pelvic floor muscle strength MSID 11.3%	104	120	140	161

11.2.3 Trial processes

For the not fulfilled feasibility criteria of this preliminary trial, the full RCT needs to consider strategies to enhance the rates. Respective potential solution strategies, changing aspects of trial design, are elaborated in Table 29, together with potential solution strategies for other important but problematic process issues. The presented solutions are suggestions only, as definitive solutions together with their assessment and evaluation (the ensuing steps in Bugge et al.'s (2013) framework) can only be determined by the research team of the full RCT for its respective research context. A number of suggested further minor modifications for the full RCT research protocol are summarised in Box 13. This includes the characteristics that were already modified during this feasibility trial and presented in Table 23.

Recording adherence might also be performed via an electronic application as was tested for dietary information by Voils et al. (2018); as positive reinforcement, notice of reached adherence could be provided to participants. Of strategies suggested to enhance adherence (Laycock, 2008, Hulley et al., 2013), improving communication and comprehension of intervention principles could be applied in a full trial by e.g. explaining training physiology in more detail; likewise, knowledge and perception of the pelvic floor are adherence facilitators (Alewijnse et al., 2007). Considering the strategies used to enhance adherence to pelvic floor muscle training identified in five trials by Dumoulin et al. (2014), a technique which might be useful in a full RCT is a smartphone-based reminder system (Kinouchi and Ohashi, 2018,

Propagator and NHS, 2018). Strategies such as red stick-up dots which were used in a trial that showed high adherence to postpartum pelvic floor muscle training for women with urinary incontinence (Chiarelli et al., 2003). Automatic e-mail reminders for the interventions were suggested by a participant. As, according to Williams et al. (2006), frequent contact with professionals and their advice, encouragement and support enhanced motivation in urinary incontinent participants, and motivation again is associated with adherence (Alewijnse et al., 2007), contact with the researcher might be intensified in a future trial. However, suggestions from trials on urinary incontinence might not be as effective in a prevention only trial.

With the modification of the experimental intervention to increasing ball weight, participants' clinical experience might change. Therefore, keeping respective questions (e.g. "Did it feel like the ball would slip out?") in a full RCT might be useful. A postintervention online questionnaire (with paper option) to survey participants' experiences with and opinion on the trial should be kept in a full RCT as this would allow for participants' anonymous feedback.

A general recommendation for a future RCT is to choose a trial setting where women know and might like to try the vaginal balls. As recruitment, for the different context, might differ in a full trial, Thoma et al.'s (2010) suggestions on how to optimise recruitment might be considered, and Donovan et al.'s (2016) Quintet Recruitment Intervention instrument can be consulted to tackle new recruitment issues that might emerge in different settings. Likewise, as a good retention rate needs to be assured (Daykin et al., 2018), retention strategies as described by Brueton et al. (2013) might be considered; methods to increase response to electronic questionnaires might also be of interest (Edwards et al., 2009).

Table 29 Identification of problems with feasibility criteria and other important process issues, and solutions suggested for a full RCT

Problem area	Suggested solution
Sample representativeness	• Apply strategies to enhance sample representativeness in terms of age, education and ethnicity/cultural background, e.g. translation services
Feasibility criterion 3: Intervention start within 4 days	• Inform participants to start the intervention within 4 days after perineometry • Do not recruit women before planned absences • Set target to intervention start within 7 days
Feasibility criterion 4b: Adequate adherence	• Inform about necessary adherence • Implement strategies to facilitate adherence, apply reminder methods • Use newly developed wider adherence criteria
Feasibility criterion 5: Final data collection within 2 weeks after intervention end	• Offer perineometry appointments on a more flexible timescale • Combine the two final appointments
Potential harm of vibrating vaginal ball use	• Consult hygienist about ball cleaning recommendations • Develop ball cleaning schedule to be applied after vulvovaginal symptoms • Refine harms data collection in a full trial • Ensure medical examination and diagnosis of vulvovaginal symptoms

Box 13 Minor process modifications for a full RCT

Keep modifications already implemented in feasibility trial (suggested):

- Amended selection criteria
- Optimised trial forms
- Reminder text message to be sent the day before the measurement appointment
- Unlimited number of adverse event screening calls until participant can be reached
- Question about adverse events at final call
- Online survey reminder at final call and per mail (instead of planned text messages), 2 and 4 weeks after being sent the link
- Optimised SPSS codes
- More than one data validation method
- Application of rule of three for harms analysis

Suggested further modifications:

- Adapt recruitment to new trial context (preferably active route with personal contact)
- Delete exclusion criterion: “Termination of participation will be recommended to a participant when there is vaginal infection”
- Amend participant information and consent form to include possiblity to perform pelvic floor muscle training if desired by participants in experimental group
- Alert potential participants to the potential burden of trial participation
- Offer home visits before noon, and also for perineometry
- Offer perineometry appointments more often and at convenient times
- Send a text message reminder before each trial appointment
- To ensure blinding, include the respective instruction in the routine text message sent before each measurement appointment; add reminder to the measurement standard
- Offer babysitting service at perineometry appointments at study site
- Enhance measurement standardisation by enhancing instruction of participants, training of assessors (including a thorough introduction to the trial), and closer result monitoring by PI
- Ensure online access to perineometry forms or perineometry values’ confidentialty at all times
- Do not let participants take pictures of their perineometry results
- Where applicable, refine data collection forms (suggestions in Appendix FF)
- Adapt content of interviews schedules and online survey
- Instruct participants not to keep the ball in the vagina for substantially longer than 30 minutes
- Consider sitting times during ball use
- Offer electronic application instead of paper sheet to document adherence
- Keep from reminders of correct training at the adverse event phone calls
- Develop ball cleaning scheme to be applied after vulvovaginal symptoms

Modifications to consider:

- Vaginal examination by the researcher before inclusion at the initial study meeting
- Enquire subjective pelvic floor perception in more detail
- Change of adherence criteria (e.g. considering training three times/week as adherent), calculation of adherence by different statistical procedures and sensitivity analysis
- Provide participants with an alarm clock
- Provide participants with soap samples
- Standardising duration of strength contractions
- Offer forms for trial staff as electronic versions

11.2.4 Trial management and resources

Lessons learnt from the experiences with this feasibility trial suggest for:

- Human management: funding will ease finding pelvic floor assessors; assessors should get a more thorough introduction into their task, and they need to be insured and risk assessed. Organising a rota for the measurement appointments is a challenging task.
- Organisational management: be aware of vaginal measurement as an intimate topic; home visits are recommended, if possible also for perineometry measurements; otherwise look for a friendly venue with a locker for study materials.
- Data management: take precautions to keep perineometry data confidential at all times, and develop a document naming/numbering system before trial start.

Further, in planning a full RCT, Tickle-Degnen's (2013), Thabane et al.'s (2010) and Orsmond and Cohn's (2015) questions on trial management (partly reproduced in Appendix CC) can serve as a guidance.

Likewise, the resources calculation of this feasibility trial, compiled for the categories facilities, staff, materials and transport, and complemented by these authors' resources questions (also partly reproduced in Appendix CC), can serve as the basis for the budget calculation for a full RCT. A definitive cost plan can only be established after PPI work for the future trial and having taken the necessary decisions about the full RCT's design, and will be based on the country, research setting and context. However, to give a rough idea about the costs of a full RCT, a preliminary cost plan is prepared, based on the resources calculation for this feasibility trial in section 7.2.

This preliminary cost plan assumes the same context as in this study. The pelvic floor centre at the AKH Vienna would be the project leading institution, its facilities and some items (shown in Table 30) are assumed to be available without extra costs and therefore are not included in the calculation. It is also assumed that the trial design is identical, with the exception of using three balls of enhancing weight per participant. The second part of Table 30 shows the items which need to be purchased, and their cost per category for minimum and maximum sample size. The prices are according to information obtained from the world wide web between June and August 2019. For the minimum number of 104 participants, this results in a sum of 66,735 € (56,852 £ on 7 December 2019); for the maximum number of 773 participants, this results in a sum of 314,800 € (268,180 £ on 7 December 2019).

Table 30 Cost plan

Items assumed to be available at research site and not included in cost calculation			
Category	**Item**		
Facilities	- Office with computer, printer and web accessibility - Room for pelvic floor measurements (lockable) - Waiting area with toilet - Lockable storage place for Peritron - Library access		
Staff	- Liability insurance for hands-on scientific staff		
Materials	- Mobile phone - Computer programmes: Office 365, SPSS, R, endnote - Subscription to web survey application software - Peritron		
Expected costs			
Category	**Item**	**Euro** ***N*** **= 104**	**Euro** ***N*** **= 773**
Staff	- Project lead - Recruitment - Organising visits and measurement appointments - Data collection - Perineometry - Data processing and analysis - Report writing	60823.00	276936.00
Materials	- Photocopies - Stationeries - Postage - Catering - Vaginal balls - Replacement balls - Lubricant for ball use - Antibacterial soap - Examination gloves - Probe condom covers - Lubricant for measurement - Exam table paper - Disinfection wipes - Spares	4258.73	29197.16
Transport	- Tickets for measurement attendance - Tickets for home study visits - 2% spares	1653.76	8668.62

In this feasibility trial, participants incurred no financial costs for trial participation as they were provided with the study materials and reimbursed their travel expenses; this is also suggested for a full trial. The nonfinancial study contribution needed from each participant is shown in Figure 14. This comprised the daily intervention time (including adherence documentation), the time for the information/consent/initial and final study visits (1.5 and 0.75 hours respectively), initial and final perineometry (travel/waiting time and 15 minutes per measurement), at least five phone calls à 5-10 minutes, and, for half of them, completing the final online questionnaire (15-20 minutes). Although, according to future design and methods decisions, these efforts might change in a full RCT, this summary provides an approximate estimate of what will be expected from future participants in terms of time resources.

Summary

The feasibility criteria that were not fulfilled in this trial and further results for trial processes and management ask for a number of modifications, with some modifications already having been implemented during the trial. Resources calculation and a preliminary cost plan for a future trial could be compiled. The clinical results of the present feasibility trial represent an initial attempt to assess the clinical outcomes of using a vibrating vaginal ball to strengthen the pelvic floor muscles after childbirth. Compared to enhanced standard care, the preliminary effect results show, although slightly contradictory, a tendency towards a larger effect in the comparison group. Preliminary harms results show fewer harms in the comparison group. Different feasibility factors render the results potentially biased. The trial also showed that the vibrating vaginal balls were well received by participants. However, in view of the adverse events encountered, and although none were serious, it must be discussed whether a future trial should be done. PPI work can help with this and with decisions on other trial characteristics.

In case a future full trial will be planned, a number of implications is suggested. These concern the trial design, participants, choice of interventions and comparisons, the most appropriate future primary outcome measure, and analysis, as well as a number of major and minor process implications. Depending on primary outcome and desired power, the approximate necessary sample size for the full RCT is estimated as between 104 and 773.

12 CONCLUSION

This concluding chapter draws the knowledge gained in this doctoral work together and outlines its contribution to existing knowledge. After an overview on the findings from the systematic review and the feasibility trial, it highlights strengths and limitations of the empirical part. The chapter and thesis end with implications for research and professional practice.

12.1 Review of thesis findings

As this PhD project consisted of a systematic review and an empirical study, the summary of the findings is presented in two sections.

12.1.1 Findings from the systematic review

The quantitative systematic review of existing scientific evidence aimed to compare the effectiveness of vaginal balls or cones for improvement of pelvic floor muscle performance and urinary continence in the postpartum period to no treatment, placebo, sham treatment or a comparison intervention. It included one RCT on weighted vaginal cones to treat urinary incontinence in the first year after birth. The review results suggest that the use of cones, when compared with standard pelvic floor care or an enforced exercise regimen, might be helpful for urinary incontinence up to two years after childbirth, but the wide effect size CIs indicate that their use might equally not be helpful. Likewise, with small effect sizes and wide CIs, a meaningful difference of pelvic floor muscle performance could not be shown between cone and standard care group (where a difference would be desired) nor between cone and enforced pelvic floor muscle exercises group (where no difference is desired[122]). Further, the validity of the ITT reanalysis for the only trial that met the inclusion criteria is limited by the high rate of withdrawals, especially in the intervention groups, the lack of participant blinding, and by sample size not being based on a power calculation. Also, the pelvic floor muscle exercise regimen used in the study does not correspond to contemporary evidence-based recommendations.

With respect to the effectiveness of vibrating vaginal balls, the review showed a dearth of respective scientific evidence as no study on the vaginal use of vibrating balls in the postpartum time could be identified. It was therefore concluded that further research was needed and the trial planned for this doctoral study was initiated.

[122] It must be stressed however that finding no difference in a superiority trial does not imply equivalence: with a nonsignificant result, no difference between groups cannot be concluded but can only not be ruled out (Barker et al., 2002, Walker and Nowacki, 2011).

12.1.2 Findings of empirical study

The feasibility trial tested the feasibility and methodology of a planned full RCT to evaluate the effectiveness of vibrating vaginal balls to strengthen the pelvic floor muscles after childbirth. It was able to fulfil its aims and objectives and to answer its research questions: It succeeded in assessing feasibility and practical issues of a future full RCT (including methodological questions and sample size), in gaining preliminary effect results and monitoring potential harms of the experimental intervention, and in exploring women's perspectives on and experiences with the interventions and the trial. A summary of the issues covered in this feasibility trial against 14 methodological issues to be evaluated in pilot and feasibility trials (Shanyinde et al., 2011, Bugge et al., 2013) is presented in Table 31.

This feasibility trial's results indicate that a full RCT to investigate the intervention of interest seems feasible. The trial processes recruitment, randomisation and data collection generally ran smoothly. Vibrating vaginal ball use was feasible in the postpartum sample although arranging the necessary 30 minutes was sometimes difficult for participants and they felt bound timewise or to the house. Also, both interventions lacked adherence as a difficulty was the need to remember and to find the time to do them alongside childcare. Two of the five feasibility criteria–on recruitment and timely preintervention measurement attendance–were fulfilled. The rates of those which were not, and this included adherence, timely intervention start and final data collection, might be enhanced with trial modifications. The clinical results with their wide CIs (which include no effect or an effect in the other direction) motivate the conduct of a larger, more definitive trial to come to a more firm conclusion on effect and harms. Suggestions to enhance trial management could be extracted from the collected information, and a preliminary cost plan could be calculated. Participants' feedback on their trial experience was mainly positive and encouraging, and they provided a number of suggestions to enhance the trial.

Table 31 Issues assessed in this feasibility trial (according to Shanyinde et al. (2011) and Bugge et al. (2013))

Methodological issues	Findings	Evidence in section
Did the feasibility study allow a sample size calculation for the main trial?	Yes	10.3.2
What factors influenced eligibility and what proportion of those approached were eligible?	24.3% of women (18/74) checked for were not fulfilling the eligibility criteria, leaving 75.7% as eligible. 86.6% of the women (116/134) at the researcher level were potentially eligible. This is unclear for 60 of them because other reasons hindered participation before eligibility was checked.	6.1.1
Was recruitment successful?	Yes in quantitative terms: • Targeted sample size recruited within time planned • Recruitment rate: 47.1% of eligible women • Declining participation: 28.7% of women at researcher level No in qualitative terms: • Recruitment path via hospital not successful • Sample not representative of target population	6.1.1
Did eligible participants consent?	At least 47.1% (95% CI [38.1, 56.1]) of approached eligible women consented, which fulfilled Feasibility criterion 1.	6.1.1
Were participants successfully randomised? Did randomisation yield equality in groups?	Yes, all were randomised. Comparability of randomised groups reached for all considered variables except for 2nd degree perineal tear and ICIQ-UI SF sum score/urinary incontinence rate.	6.3 9.2.1
Were blinding procedures adequate?	Mostly: Only two participants revealed group allocation to assessors.	7.1.3
Did participants adhere to the intervention?	Not adequately: Depending on the calculation method, only 47.2% of participants (25/53, 95% CI [33.8, 60.6]) minimum or 60.4% (32/53, 95% CI [47.2, 73.6])	7.2.2

Methodological issues	Findings	Evidence in section
	maximum adhered to their intervention. Cave: The appropriate level of adherence is uncertain.	
Were the interventions acceptable to the participants?	Yes, as expressed in final interviews and by adherence rates.	7.2
Was it possible to calculate cost and duration of study?	Yes	8.2, 6.4
Were outcome assessments completed?	Yes	7.1.3
Were outcomes measured those that were the most appropriate outcomes?	Yes	10.3.1
Was retention to the study good?	Yes: Only one participant withdrew, retention rate was 98.2% (54/55).	6.4
Were the logistics of running a multicenter trial assessed?	No	n.a.
Did all components of the protocol work together?	Yes	10.1

On the basis of the information gathered, a full RCT can be designed. Thereby, the present trial protocol would need to be modified by amending elements of trial design and operationalisation to address the issues identified as suboptimal and needing improvement. However, before finalising a protocol for a full RCT, PPI participants need to comment on whether a full RCT is justified and desired from their point of view as there is an important possibility of adverse events. PPI participants also need to be consulted on whether they think the research question with the primary outcome pelvic floor muscle strength is of (enough) interest to women, and to review the modified research protocol.

12.2 Strengths and limitations of this feasibility trial

Strengths of the present feasibility trial as well as its limitations are considered in the following two sections. In each section, this is done for the trial's internal validity and reliability and for its external validity (generalisability), and in each case for both the feasibility and the clinical outcomes. Although the clinical outcomes are part of the feasibility outcomes, they are considered in separate sections to stress their preliminary nature.

12.2.1 Strengths

This feasibility trial is, to the author's knowledge, the first study to examine the use of vibrating vaginal balls after childbirth. It fulfilled its stated aims and objectives and provides evidence on the feasibility of a full RCT to determine the effectiveness of vibrating vaginal balls for pelvic floor muscle strengthening after childbirth; this includes women's opinion on and experiences with the device and the trial. The trial has progressed existing knowledge with respect to, e.g., intervention acceptability and feasibility, factors affecting adherence, or issues that matter to women. It also provides preliminary scientific evidence on effect and harms of postpartum ball use.

Internal validity and reliability

Feasibility. Different methodological features enhance the internal validity of this trial's feasibility results. The trial had clear feasibility aims and predefined progression criteria. A scientifically justified sample size calculation determined the appropriate number of participants who were specified via selection criteria. The key elements ensuring validity and reliability of the future RCT–random group allocation, allocation concealment and masked assessment–were applied and tested. Relevant parts of the trial protocol were informed and enhanced by PPI participants' input, as forms developed for this feasibility trial were submitted to a PPI process and optimised during the trial. Outcome data were almost complete, and all were reported. As appropriate for a feasibility trial, it was refrained from significance and hypothesis testing. Although *p* values are available from an effect size

calculation by an mITT analysis, their “nonsignificance” in theoretical and practical terms is clearly pointed out. Sufficient data could be obtained to inform a sample size calculation for a future full trial. The experimental feasibility results were complemented by a survey that allowed exploration and consideration of participants’ experiences and opinion and which can contribute to optimally plan a full RCT. Reliability of the feasibility results was enhanced by data entry error checking through four recommended validation strategies. Risks of bias were assessed. Minor protocol modifications, such as a more comprehensive feasibility analysis, unplanned interviews with the second half of the participants, or the deletion of planned effect sensitivity analyses, strengthened the trial and improved its integrity.

Clinical results. Internal validity and reliability of effect and harms results was enhanced by the key elements ensuring validity and reliability in a future RCT–a valid random group allocation, allocation concealment (by sequentially numbered, opaque and sealed envelopes), and blinded assessment. Even if the participants and CO as PI were aware of group allocation, allocation could in almost all cases be concealed from the blinded outcome assessors. The comparison intervention pelvic floor muscle training was based on exercise science principles, and adherence was registered for both interventions. The use of two methods of data collection–PROs and technical measurement–in the same study and the same research approach enhanced outcome information on pelvic floor muscle performance by methodological triangulation; this was supported by qualitative data from open-ended interview questions. The pelvic floor questionnaire contained validated items and the ICIQ-UI SF as a validated questionnaire. The perineometric measurements for muscle strength and, to a smaller degree, for vaginal resting pressure are, as performed with care, valid and reliable. Strategies to optimise perineometry accuracy[123] and reliability were a measurement standard, measurement repetitions for contraction strength, training the observers, and keeping them blinded. Reliability of the clinical results was enhanced by data entry error checking through four recommended validation strategies. There was low attrition, mITT analysis allowed consideration of all participants as randomised, a PP analysis supported the findings of the mITT analysis for the potential two future primary clinical outcomes, and all (almost complete) outcomes were reported. The preliminary harms analysis was strengthened by keeping to international guidelines on harms reporting in trials and by application of the rule of three. A risk of bias assessment was performed for effect and harms results.

External validity

The characteristics named as enhancing internal validity also apply to external validity. The inclusion of participant reported pelvic floor outcomes should increase the relevance for

[123] The degree to wich a measurement represents the true value (Hulley et al., 2013).

(potential) ball users. Further, external validity of the results was strengthened by PPI participants' input from preparatory PPI work in the UK and Austria.

12.2.2 Limitations

There are several limitations to this feasibility trial, stemming from design and planned methodology and from study realisation.

Internal validity and reliability

Feasibility. The feasibility trial had no pilot stage itself and therefore lacked such a stage's corrective function. Feasibility assessments may be misleading by the small number of highly motivated, nonrepresentative participants sampled by a nonrandom strategy in a single city in Austria. By only using one ball weight, the intervention was not applied in its potential fullest form. Part of the measurement instruments were self-designed and not formally tested for validity and reliability before the trial, and the descriptive part of the feasibility data collection partly also relied on informal processes. Likewise, data collection processes were not tested or formally evaluated before the trial. More detailed answering categories for the interview questions would have enabled more systematic probing and a more refined information collection. CO as interviewer may have influenced the participants' responses towards social desiredness, and although answers were jotted down during the interview with great care, content may have been missed or misunderstandings may have happened. Qualitative data transcription and analysis is an interpretative process. As no funding was available, there was no independent data analysis with results cross checking by another researcher. There are various sources of potential bias for the feasibility results.

Clinical results. Most of the named limitations to the feasibility results also apply to the internal validity and reliability of this feasibility trial's potential to determine intervention effectiveness and harms. Sources of potential imprecision are the small sample and interobserver variability by three blinded assessors. The small intervention contrast (two active interventions) is biasing towards no effect as it is easier to find an effect when the intervention contrast is larger (Woodley et al., 2017). Adherence was poor; however, the level of adherence necessary for an effect is unclear for both interventions. The manometric endurance measurements' reliability is uncertain: There is no reliability study available that researched the form of operationalisation used in this trial, and the values' accuracy and reliability appear limited from screening the result values and assessors' feedback. The effect results are slightly contradictory, and the confidence in the estimates is limited as the true effects may be substantially different from the effect estimates. As some refinement in effectiveness and harms data collection forms and processes was needed during trial execution, a few outcome data are missing. The lack of a pilot phase led to suboptimal

adverse events screening, e.g. for the outcome recurrence. A few data relied on maternal recall, and adverse events were self-reported; no medical diagnosis (potentially of candidiasis) was obtained for most vulvovaginal symptoms. Confounding, moderating and mediating variables were not considered in the effect analysis, and missing data were not imputed. There is a high risk for selection bias, and an at least medium risk for performance and detection bias.

External validity

All named threats to internal validity also apply to external validity which is limited by further factors.

Feasibility. Findings of small studies in a single centre are not representative. The small sample size, nonrandom sampling strategy and nonrepresentativenss of the sample make it difficult to generalise the feasibility statement for a full trial to the target population of all childbearing women, neither in Austria nor in other countries. The results might also differ with participants who are past six months post partum. As the use of vibrating vaginal balls is culturally influenced, there might be low generalisability of recruitment and feasibility of ball use results to other than the included ethnic groups or to countries where such devices are less well known or appreciated. As a PhD project, the study might express the feasibility of doing a full trial on this topic in a PhD context which might not be applicable to other contexts. With modifications in the research protocol, such as applying the experimental intervention in its potential fullest form, trial feasibility might change. The cost plan, as calculated on the basis of the same trial context, is only tentative and not applicable to other research settings or a modified trial design (the only calculated modificaton being three balls of enhancing weight per participant).

Clinical results. The method of ball use researched was hold and vibration use only. The experimental intervention dose can be considered low as the intervention was not applied with the heaviest ball a participant could retain or increasing ball weight. The dose of the comparison intervention was higher than that in current Austrian routine postpartum pelvic floor muscle recommendations. Results might differ in women who were not included by the selection criteria and who are e.g. past six months post partum, in a different socioeconomic situation, or after Caesarean section. As the trial has a predominantly explanatory (and not pragmatic) design, its external validity holds for the question of efficacy rather than of effectiveness. Comparison of the clinical results with other studies is restricted because of the results' feasibility nature, because there are, as closest similar studies, only few trials on cones examining pelvic floor muscle performance post partum, and because measurement by different devices with different vaginal probe sizes cannot be compared (Bø et al., 2005).

12.3 Implications of thesis findings

The implications section unites the lessons learnt after reflecting on this feasibility trial's results. It outlines the novel contribution of this doctoral research to knowledge and maternal pelvic floor health. As the feasibility trial was performed with the aim to prepare a future full RCT, implications for research are specified first. Thereafter, implications for the use of vibrating vaginal balls to strengthen the pelvic floor muscles after childbirth, gained from the systematic review and the empirical part of the PhD, are stated.

12.3.1 Implications for research

The knowledge gained in this thesis supports further investigation into the topic. It does so by the result of the systematic review which suggests that the available scientific evidence base for the experimental intervention is weak, and strong evidence yet to be created. Likewise, the results of the feasibility trial (as expected) did not resolve the existing equipoise between (enhanced) standard pelvic floor muscle exercises and the use of vibrating vaginal balls to strengthen the pelvic floor muscles post partum. They point to the feasibility of a future definitive RCT although there is potential bias and remaining uncertainty about feasibility.

Based on the experiences gained in this feasibiltiy trial, necessary or potential alterations for a full scale trial protocol are suggested to enhance its methodology, validity and reliability. This concerns e.g. the choice of interventions to be compared, offering perineometry appointments on a more flexible timescale, or measures to improve adherence. However, with a view on the rate of adverse events (in particular discomfort and vulvovaginal symptoms/infection) accompanying the experimental intervention, PPI work is needed to find out whether women would be willing to accept the potential adverse effects in relation to the potential beneficial effect of the experimental intervention. Depending on the selected primary outcome, MSID, and desired power, the estimated approximate sample size for a full RCT lies between 104 and 773. Although a cost plan for a full RCT could be compiled, this must be seen as tentative as other research settings and research protocol changes may lead to different calculation results. Only a potential funder can decide whether time and money needed for the full trial are economically justifiable.

By submitting this feasibility trial's findings with a funding application for a future RCT to support the claim of feasibility, funding chances might be increased. One limitation however is that the conclusion about feasibility can only be drawn for the specific feasibility trial setting, making generalisation to other locations and settings difficult. Therefore, the recommended next step is to conduct an internal pilot trial according to a modified research protocol. Potentially, the clinical results could contribute to a future meta-analysis in this field whereby its important limits must be acknowledged. An interesting statistical calculation

might be to reanalyse the clinical data within a Bayesian framework (adding prior skeptical and enthusiastic belief) to assess whether to perform a confirmatory full RCT (Parmar et al., 1996).

12.3.2 Implications for vibrating vaginal ball use in the postpartum time

Even when "bravely" accepting the limitations of a pilot study (Leon et al., 2011, p. 628), meaning that no valid conclusions can be drawn about intervention effect and harms, preliminary implications for the use of vibrating vaginal balls in the postpartum time can be derived. This shall help inform health professionals' clinical practice and be of use to information-seeking women. Interest in this work has already been noticed, e.g. in Herman & Wallace (2018) or by e-mail communication received.

At the end of this thesis, the question about vibrating vaginal balls' (and similar devices') effectiveness to enhance pelvic floor muscle performance post partum remains unresolved. There are different methods of ball use, and the dosage of use is unclear. Issues identified regarding the feasibility of 30 minutes of hold and vibration use were the need to think about ball use and finding the necessary time. In the sample, the PROs showed a tendency of vibrating vaginal ball use (one ball of 28 g) being inferior to pelvic floor muscle training with respect to enhancing pelvic floor muscle strength and symptoms. For pelvic floor muscle resting pressure and contraction strength, the perineometry results favoured the vaginal balls. It was found that there is potential for harm, the main adverse events being discomfort and vulvovaginal symptoms/infection. Even if results for effect and harms were gained in this study, these are preliminary and not fully valid as they could originate from a random effect due to the small sample size, from the nonadjusted statistical analysis or from the discussed risks of bias.

This doctoral work showed that the scientific evidence base to date is not firm enough to recommend vibrating vaginal balls in postpartum care, neither to improve the pelvic floor muscles nor to enhance urinary incontinence (or other) symptoms; however, neither is the evidence base strong enough to discourage from the use of this available device for the named purposes. Also unresolved remains the question about other purposes of ball use, such as enhancing pelvic floor muscle awareness or sexual sensations. As the clinical results are preliminary and limited by the discussed factors, no change to present (Austrian) maternal healthcare policy (usual care) is recommended. The information gained in this thesis, however, can be communicated in professional education which would then equip professionals to better respond to any questions from clients.

Summary

Findings from this doctoral work are available from a systematic review and an empirical study. The results from the systematic review showed that carrying weighted vaginal cones for urinary incontinence after childbirth might be helpful or not when compared to the included trial's enforced pelvic floor muscle exercise regimen or to usual care. No research could be identified on vibrating vaginal balls to enhance urinary incontinence or improve pelvic floor muscle performance post partum.

The empirical part of this doctoral work consisted in a randomised controlled feasibility trial to prepare a full RCT to determine vibrating vaginal balls' ability to improve pelvic floor muscle performance after childbirth. This feasibility RCT provided valuable key insights into trial processes, management, resources, clinical results, and participants' experiences and opinion to inform a full RCT. It indicates that a full RCT to investigate the effect of interest seems feasible and worthwhile, at the same time pointing to areas needing improvement. Suggested modifications for a full trial refer to design issues, the interventions, outcome measures, and trial forms. On the basis of the information collected about harms, PPI work is recommended to gather women's opinion on a future full RCT; this PPI work shall also clarify women's general interest in the proposed study before an internal pilot trial of a full RCT will be designed.

The trial followed robust procedures which enhanced internal and external validity of the feasibility (including the clinical) results. However, several limitations render this feasibility trial's results potentially biased, e.g. the restriction of the research setting to a single Austrian city, the nonrepresentativeness of the small sample, the applied simple version of the experimental intervention, or the lack of adherence. The generalisability of this trial's feasibility results therefore is questionable, and even if they seem promising, it cannot be concluded with confidence whether a full RCT will be feasible with a modified protocol and in another research setting.

Regarding the use of vibrating vaginal balls in the postpartum period, a more informed decision than hitherto is possible. From this work, it can be concluded that to date there is not enough scientific evidence to justify any form of use of this device to improve pelvic floor muscle performance and/or urinary continence after childbirth. However, neither can women be discouraged from the use of vibrating vaginal balls on the basis of the available scientific evidence. Potential adverse effects need to be taken into account when considering ball use.

REFERENCES

1 Cor 13: 2, Holy Bible. New international version. Available at: http://biblehub.com/1_corinthians/13-2.htm (Accessed: 26 June 2018).

Abbott, J.H. (2014) 'The distinction between randomized clinical trials (RCTs) and preliminary feasibility and pilot studies: what they are and are not', *The Journal of Orthopaedic and Sports Physical Therapy*, 44(8), pp. 555-558.

Abdel Karim Ruiz, M., Gallego Molina, J., and Gavilán Díaz, M. (2014) 'Äreas vaginais: função terapêutica', *Cultura de los Cuidados*, 18(40), pp. 93-98. doi: 10.7184/cuid.2014.40.12.

Abdool, Z., Shek, K.L., and Dietz, H.P. (2009a) 'The effect of levator avulsion on hiatal dimension and function', *American Journal of Obstetrics & Gynecology*, 201(1), pp. 89.e1-89.e5.

Abdool, Z., Thakar, R., and Sultan, A.H. (2009b) 'Postpartum female sexual function', *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 145(2), pp. 133-137.

Abrams, P., Avery, K., Gardener, N., and Donovan, J. (2006) 'The International Consultation on Incontinence Modular Questionnaire: www.iciq.net', *The Journal of Urology*, 175(3 Pt 1), pp. 1063-1066.

Abrams, P., Cardozo, L., Wagg, A., and Wein, A. (2017) *Incontinence*. 6th edn. No place: ICUD ICS.

Afshari, P., Dabagh, F., Iravani, M., and Abedi, P. (2017) 'Comparison of pelvic floor muscle strength in nulliparous women and those with normal vaginal delivery and cesarean section', *International Urogynecology Journal*, 28(8), pp. 1171-1175.

Ahlund, S., Nordgren, B., Wilander, E.L., Wiklund, I., and Friden, C. (2013) 'Is home-based pelvic floor muscle training effective in treatment of urinary incontinence after birth in primiparous women? A randomized controlled trial', *Acta Obstetricia et Gynecologica Scandinavica*, 92(8), pp. 909-915.

Alewijnse, D., Mesters, I., Metsemakers, J.F.M., and van den Borne, B. (2007) 'Strategies to enhance adherence and reduce drop out in conservative treatment', in Bø, K., Berghmans, B., Mørkved, S. & van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* Edinburgh: Churchill Livingstone, pp. 133-146.

Allgemeines Krankenhaus der Stadt Wien (2017) E-mail to Claudia Oblasser. 29 September.

Allto Consulting (2018) *Confidence interval calculator for proportions*. Available at: https://www.allto.co.uk/tools/statistic-calculators/confidence-interval-for-proportions-calculator/ (Accessed: 23 July 2018).

Amaro, J.L., Moreira, E.C.H., Gameiro, M.O., and Padovani, C.R. (2005) 'Pelvic floor muscle evaluation in incontinent patients', *International Urogynecology Journal*, 16(5), pp. 352-354.

American College of Sports Medicine (2002) 'American College of Sports Medicine position stand. Progression models in resistance training for healthy adults', *Medicine and Science in Sports and Exercise*, 34(2), pp. 364-380.

American College of Sports Medicine (2009) 'American College of Sports Medicine position stand. Progression models in resistance training for healthy adults', *Medicine and Science in Sports and Exercise*, 41(3), pp. 687-708.

American Psychological Association (2010) *Publication manual of the American Psychological Association*. 6th edn. Washington, DC: American Psychological Association.

Amorelie (no date) *Vibro-Eier*. Available at: https://www.amorelie.de/sexspielzeug/vibratoren/vibro-eier/ (Accessed: 26 June 2018).

Anderson, M.R., Anderson, M.R., Klink, K., and Cohrssen, A. (2004) 'Evaluation of vaginal complaints', *JAMA*, 291(11), pp. 1368-1379.

Angelo, P.H., Varella, L.R.D., de Oliveira, M.C.E., Matias, M.G.L., de Azevedo, M.A.R., de Almeida, L.M., de Azevedo, P.R.M., and Micussi, M.T. (2017) 'A manometry classification to assess pelvic floor muscle function in women', *PLoS ONE*, 12(10), p. e0187045. doi: 10.1371/journal.pone.0187045.

Antara Training (2018) *Antara®*. Available at: http://www.antara-training.ch/home.aspx (Accessed: 26 June 2018).

Antara Training Österreich (2018) *Das Bewegungs[sic] und Trainingkonzept*. Available at: http://antara-oesterreich.at/antara/das-bewegungs-und-trainingkonzept/ (Accessed: 26 June 2018).

Arbeitsgemeinschaft für Urogynäkologie und rekonstruktive Beckenbodenchirurgie Österreich (2014) *Leitlinie zum Management von Dammrissen III. und IV. Grades nach vaginaler Geburt*. Available at: http://www.urogyn.at/wp-content/uploads/2015/10/LLDR.pdf (Accessed: 29 June 2018).

Arnold, D.M., Burns, K.E., Adhikari, N.K., Kho, M.E., Meade, M.O., and Cook, D.J. (2009) 'The design and interpretation of pilot trials in clinical research in critical care', *Critical Care Medicine*, 37(1 Suppl), pp. S69-S74.

Arvonen, T., Fianu-Jonasson, A., and Tyni-Lenné, R. (2001) 'Effectiveness of two conservative modes of physical therapy in women with urinary stress incontinence', *Neurourology and Urodynamics*, 20(5), pp. 591-599.

Arvonen, T., Jonasson, A., and Tyni-Lenné, R. (2002) 'A clinical trial comparing conventional pelvic floor training and training with vaginal balls in women with stress urinary incontinence - a pilot study', *Nordisk Fysioterapi*, 6(1), pp. 41-46.

Ashton-Miller, J., Howard, D., and Delancey, J. (2001) 'The functional anatomy of the female pelvic floor and stress continence control system', *Scandinavian Journal of Urology and Nephrology*, 35(2), pp. 1-7.

Ashton-Miller, J.A. and Delancey, J.O. (2009) 'On the biomechanics of vaginal birth and common sequelae', *Annual Review of Biomedical Engineering*, 11(1), pp. 163-176.

Ashton-Miller, J.A. and DeLancey, J.O.L. (2015) 'Functional anatomy of the female pelvic floor', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone, pp. 19-34.

Ashworth, P.D. and Hagan, M.T. (1993) 'Some social consequences of non-compliance with pelvic floor exercises', *Physiotherapy*, 79(7), pp. 465-471.

Association of Chartered Physiotherapists in Women's Health (2013a) *Fit and safe to exercise in the childbearing year (advice for physiotherapists and other health professionals).* Available at: http://pogp.csp.org.uk/publications (Accessed: 21 June 2016).

Association of Chartered Physiotherapists in Women's Health (2013b) *Fit for the future: essential advice and exercises following childbirth.* Available at: http://pogp.csp.org.uk/publications/fit-future (Accessed: 31 July 2018).

AusVet (2018) *Calculate confidence limits for a sample proportion*. Available at: http://epitools.ausvet.com.au/content.php?page=CIProportion (Accessed: 26 June 2018).

Avery, K., Donovan, J., Peters, T.J., Shaw, C., Gotoh, M., and Abrams, P. (2004) 'ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence', *Neurourology and Urodynamics*, 23(4), pp. 322-330.

Baessler, K. and Bell, B.E. (2008) 'Alternative methods to pelvic floor muscle awareness and training', in Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A. & Stanton, S.L. (eds.) *Pelvic floor re-education: principles and practice.* 2nd edn. London: Springer, pp. 208-212.

Baessler, K. and Junginger, B. (2011) 'Beckenboden-Fragebogen für Frauen: Validierung eines Instrumentes mit posttherapeutischem Modul zur Evaluation von Symptomen, Leidensdruck, Lebensqualität, Verbesserung und Zufriedenheit', *Aktuelle Urologie*, 42(5), pp. 316-322.

Baessler, K. and Kempkensteffen, C. (2009) 'Validierung eines umfassenden Beckenboden-Fragebogens für Klinik, Praxis und Forschung', *Gynäkologisch-geburtshilfliche Rundschau*, 49(4), pp. 299-307.

Baessler, K. and Schüssler, B. (2008) 'The effects of pregnancy and childbirth on the pelvic floor', in Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A. & Stanton, S.L. (eds.) *Pelvic floor re-education: principles and practice.* 2nd edn. London: Springer, pp. 36-48.

Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A., and Stanton, S.L. (2008) *Pelvic floor re-education: principles and practice*. 2nd edn. London: Springer.

Bagley, H.J., Short, H., Harman, N.L., Hickey, H.R., Gamble, C.L., Woolfall, K., Young, B., and Williamson, P.R. (2016) 'A patient and public involvement (PPI) toolkit for meaningful and flexible involvement in clinical trials–a work in progress', *Research Involvement and Engagement*, 2(1), p. 15. doi: 10.1186/s40900-016-0029-8.

Baird, D.T., McNeilly, A.S., Sawers, R.S., and Sharpe, R.M. (1979) 'Failure of estrogen-induced discharge of luteinizing hormone in lactating women', *The Journal of Clinical Endocrinology & Metabolism*, 49(4), pp. 500-506.

Baracho, S.M., Barbosa da Silva, L., Baracho, E., Lopes da Silva Filho, A., Sampaio, R.F., and Mello de Figueiredo, E. (2012) 'Pelvic floor muscle strength predicts stress urinary incontinence in primiparous women after vaginal delivery', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 23(7), pp. 899-906.

Barber, M.D., Bremer, R.E., Thor, K.B., Dolber, P.C., Kuehl, T.J., and Coates, K.W. (2002) 'Innervation of the female levator ani muscles', *American Journal of Obstetrics & Gynecology*, 187(1), pp. 64-71.

Barber, M.D., Walters, M.D., and Bump, R.C. (2005) 'Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7)', *American Journal of Obstetrics & Gynecology*, 193(1), pp. 103-113.

Barbosa, P.B., Franco, M.M., Souza, F.d.O., Antônio, F.I., Montezuma, T., and Ferreira, C.H.J. (2009) 'Comparison between measurements obtained with three different perineometers', *Clinics*, 64(6), pp. 527-533.

Barclay, L., Everitt, L., Rogan, F., Schmied, V., and Wyllie, A. (1997) 'Becoming a mother - an analysis of women's experience of early motherhood', *Journal of Advanced Nursing*, 25(4), pp. 719-728.

Barker, L.E., Luman, E.T., McCauley, M.M., and Chu, S.Y. (2002) 'Assessing equivalence: an alternative to the use of difference tests for measuring disparities in vaccination coverage', *American Journal of Epidemiology*, 156(11), pp. 1056-1061.

Baron, R.M. and Kenny, D.A. (1986) 'The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations', *Journal of Personality and Social Psychology*, 51(6), pp. 1173-1182.

Barrett, G., Pendry, E., Peacock, J., Victor, C., Thakar, R., and Manyonda, I. (2000) 'Women's sexual health after childbirth', *BJOG: An International Journal of Obstetrics & Gynaecology*, 107(2), pp. 186-195.

Basson, R., Berman, J., Burnett, A., Derogatis, L., Ferguson, D., Fourcroy, J., Goldstein, I., Graziottin, A., Heiman, J., and Laan, E. (2001) 'Report of the international consensus development conference on female sexual dysfunction: definitions and classifications', *Journal of Sex & Marital Therapy*, 27(2), pp. 83-94.

Baumann, B. (2012) *Aktuelle Praxis der Diagnostik und Therapie von Beckenbodenerkrankungen bei niedergelassenen Gynäkologen.* PhD thesis. Universität Charité Berlin. Available at: http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_000000010414/Dissertation_BrittaBaumann.pdf (Accessed: 26 June 2018).

Baxter, S. (1974) 'Labour and orgasm in primiparae', *Journal of Psychosomatic Research*, 18(3), pp. 209-216.

Bayer AG, Zonagen Inc., and Target Health Inc. (2000) *Female Sexual Function Index (FSFI)*. Available at: http://www.fsfiquestionnaire.com/ (Accessed: 26 June 2018).

Baytur, Y.B., Serter, S., Tarhan, S., Uyar, Y., Inceboz, U., and Pabuscu, Y. (2007) 'Pelvic floor function and anatomy after childbirth', *Journal of Reproductive Medicine*, 52(7), pp. 604-610.

Beard, D., Rees, J., Rombach, I., Cooper, C., Cook, J., Merritt, N., Gray, A., Gwilym, S., Judge, A., and Savulescu, J. (2015) 'The CSAW Study (Can Shoulder Arthroscopy Work?)–a placebo-controlled surgical intervention trial assessing the clinical and cost effectiveness of arthroscopic subacromial decompression for shoulder pain: study protocol for a randomised controlled trial', *Trials*, 16(1), p. 210. doi: 10.1186/s13063-015-0725-y.

Berghmans, B. and Bols, E. (2015) 'Conservative interventions for treatment of AI', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* Edinburgh: Churchill Livingstone, pp. 319-333.

Beyond Basics Physical Therapy (2018) *Anatomy of the pelvic floor*. Available at: http://www.beyondbasicsphysicaltherapy.com/pelvic-floor-anatomy (Accessed: 23 July 2018).

Bhargava, K. and Bhargava, D. (2007) 'Evidence based health care: a scientific approach to health care', *Sultan Qaboos University Medical Journal*, 7(2), pp. 105-107.

Bick, D., MacArthur, C., Knowles, H., and Winter, H. (2004) *Evidenzbasierte Wochenbettbetreuung und -pflege: Praxishandbuch für Hebammen und Pflegende*. Bern: Huber.

Bidonde, J., Busch, A.J., van der Spuy, I., Tupper, S., Kim, S.Y., and Boden, C. (2017) 'Whole body vibration exercise training for fibromyalgia', *Cochrane Database of Systematic Reviews*, (9). doi: 10.1002/14651858.CD011755.pub2.

Billingham, S.A., Whitehead, A.L., and Julious, S.A. (2013) 'An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database', *BMC Medical Research Methodology*, 13, p. 104. doi: 10.1186/1471-2288-13-104.

BIPA (2018) *Liebeskugeln*. Available at: http://www.bipa.at/liebeskugeln/toys/love?wt_ad=36960227128_206900578789&wt_kw=e_36960227128_bipa%20liebeskugeln&gclid=EAIaIQobChMIj6eh0bab2AIVC40bCh3xGgZ_EAAYASAAEgKYZfD_BwE (Accessed: 26 June 2018).

Birk, A. (2012) 'Frauen- und Paarkurs mit Schwerpunkt Stillförderung', in Deutscher Hebammenverband (ed.) *Geburtsvorbereitung: Kurskonzepte zum Kombinieren.* 2nd edn. Stuttgart: Hippokrates, pp. 280-364.

Bland, J.M. (2009) 'The tyranny of power: is there a better way to calculate sample size?', *BMJ*, 339(14 Nov), p. b3985.

Bø, K. (1995a) 'Pelvic floor muscle exercise for the treatment of stress urinary incontinence: an exercise physiology perspective', *International Urogynecology Journal*, 6(5), pp. 282-291.

Bø, K. (1995b) 'Vaginal weight cones. Theoretical framework, effect on pelvic floor muscle strength and female stress urinary incontinence', *Acta Obstetricia et Gynecologica Scandinavica*, 74(2), pp. 87-92.

Bø, K. (2003) 'Pelvic floor muscle strength and response to pelvic floor muscle training for stress urinary incontinence', *Neurourology and Urodynamcis*, 22(7), pp. 654-658.

Bø, K. (2004) 'Pelvic floor muscle training is effective in treatment of female stress urinary incontinence, but how does it work?', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 15(2), pp. 76-84.

Bø, K. (2015a) 'Introduction', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone, pp. 44-46.

Bø, K. (2015b) 'Pelvic floor muscle training for SUI', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone, pp. 162-178.

Bø, K. (2015c) 'Vaginal squeeze pressure measurement', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone, pp. 61-66.

Bø, K. and Aschehoug, A. (2015) 'Strength training', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone, pp. 117-130.

Bø, K., Berghmans, B., Mørkved, S., and Van Kampen, M. (eds.) (2015) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone.

Bø, K. and Finckenhagen, H.B. (2003) 'Is there any difference in measurement of pelvic floor muscle strength in supine and standing position?', *Acta Obstetricia et Gynecologica Scandinavica*, 82(12), pp. 1120-1124.

Bø, K. and Frawley, H. (2015) 'Pelvic floor muscle training in prevention and treatment of pelvic organ prolapse', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone, pp. 234-242.

Bø, K., Hagen, R.H., Kvarstein, B., Jørgensen, J., Larsen, S., and Burgio, K.L. (1990a) 'Pelvic floor muscle exercise for the treatment of female stress urinary incontinence: III. Effects of two different degrees of pelvic floor muscle exercises', *Neurourology and Urodynamics*, 9(5), pp. 489-502.

Bø, K. and Herbert, R.D. (2009) 'When and how should new therapies become routine clinical practice?', *Physiotherapy*, 95(1), pp. 51-57.

Bø, K. and Herbert, R.D. (2013) 'There is not yet strong evidence that exercise regimens other than pelvic floor muscle training can reduce stress urinary incontinence in women: a systematic review', *Journal of Physiotherapy*, 59(3), pp. 159-168.

Bø, K., Hilde, G., Jensen, J.S., Siafarikas, F., and Engh, M.E. (2013) 'Too tight to give birth? Assessment of pelvic floor muscle function in 277 nulliparous pregnant women', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 24(12), pp. 2065-2070.

Bø, K., Kvarstein, B., Hagen, R., and Larsen, S. (1990b) 'Pelvic floor muscle exercise for the treatment of female stress urinary incontinence: I. Reliability of vaginal pressure measurements of pelvic floor muscle strength', *Neurourology and Urodynamics*, 9(5), pp. 471-477.

Bø, K., Kvarstein, B., Hagen, R.R., and Larsen, S. (1990c) 'Pelvic floor muscle exercise for the treatment of female stress urinary incontinence: II. Validity of vaginal pressure measurements of pelvic floor muscle strength and the necessity of supplementary methods for control of correct contraction', *Neurourology and Urodynamics*, 9(5), pp. 479-487.

Bø, K. and Larsen, S. (1992) 'Pelvic floor muscle exercise for the treatment of female stress urinary incontinence: Classification and characterization of responders', *Neurourology and Urodynamics*, 11(5), pp. 497-507.

Bø, K., Owe, K.M., and Nystad, W. (2007) 'Which women do pelvic floor muscle exercises six months' postpartum?', *American Journal of Obstetrics & Gynecology*, 197(1), pp. 49-51.

Bø, K., Raastad, R., and Finckenhagen, H.B. (2005) 'Does the size of the vaginal probe affect measurement of pelvic floor muscle strength?', *Acta Obstetricia et Gynecologica Scandinavica*, 84(2), pp. 129-133.

Bø, K. and Sherburn, M. (2005) 'Evaluation of female pelvic-floor muscle function and strength', *Physical Therapy*, 85(3), pp. 269-282.

Bø, K., Talseth, T., and Holme, I. (1999) 'Single blind, randomised controlled trial of pelvic floor exercises, electrical stimulation, vaginal cones, and no treatment in management of genuine stress incontinence in women', *BMJ*, 318(7182), pp. 487-493.

Boeije, H.R., Drabble, S.J., and O'Cathain, A. (2015) 'Methodological challenges of mixed methods intervention evaluations', *Methodology*, 11(4), pp. 119-125.

Bogle, B. (2006) 'Sexual health promotion in midwifery practice', in Bowden, J. & Manning, V. (eds.) *Health promotion in midwifery: principles and practice.* 2nd edn. London: Hodder Arnold, pp. 111-126.

Bonnar, J., Franklin, M., Nott, P., and McNeilly, A. (1975) 'Effect of breast-feeding on pituitary-ovarian function after childbirth', *British Medical Journal*, 4(5988), pp. 82-84.

Boote, J., Baird, W., and Beecroft, C. (2010) 'Public involvement at the design stage of primary health research: a narrative review of case examples', *Health Policy*, 95(1), pp. 10-23.

Boote, J., Barber, R., and Cooper, C. (2006) 'Principles and indicators of successful consumer involvement in NHS research: results of a Delphi study and subgroup analysis', *Health Policy*, 75(3), pp. 280-297.

Boutron, I., Moher, D., Altman, D.G., Schulz, K.F., and Ravaud, P. (2008) 'Methods and processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments', *Annals of Internal Medicine*, 148(4), pp. W60-W66.

Bowen, D.J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C.P., Squiers, L., Fabrizio, C., and Fernandez, M. (2009) 'How we design feasibility studies', *American Journal of Preventive Medicine*, 36(5), pp. 452-457.

Bowers, D., House, A., Owens, D., and Bewick, B. (2013) *Understanding clinical papers*. Chichester: John Wiley & Sons.

Boyle, R., Hay-Smith, E.J., Cody, J.D., and Mørkved, S. (2012) 'Pelvic floor muscle training for prevention and treatment of urinary and faecal incontinence in antenatal and postnatal women', *Cochrane Database of Systematic Reviews*, (10). doi: 10.1002/14651858.CD007471.pub2.

Braekken, I.H., Majida, M., Engh, M.E., and Bo, K. (2010) 'Morphological changes after pelvic floor muscle training measured by 3-dimensional ultrasonography: a randomized controlled trial', *Obstetrics & Gynecology*, 115(2 Pt 1), pp. 317-324.

Brækken, I.H., Majida, M., Engh, M.E., and Bø, K. (2014) 'Are pelvic floor muscle thickness and size of levator hiatus associated with pelvic floor muscle strength, endurance and vaginal resting pressure in women with pelvic organ prolapse stages I–III? A cross sectional 3D ultrasound study', *Neurourology and Urodynamics*, 33(1), pp. 115-120.

Braunholtz, D.A., Edwards, S.J., and Lilford, R.J. (2001) 'Are randomized clinical trials good for us (in the short term)? Evidence for a "trial effect"', *Journal of Clinical Epidemiology*, 54(3), pp. 217-224.

Brayshaw, E. (2003) *Exercises in pregnancy and childbirth: a practical guide for educators*. New York, NY: Books for Midwives.

Bristol Urological Institute (2014a) *ICIQ structure*. Available at: http://www.iciq.net/structure.html (Accessed: 29 June 2018).

Bristol Urological Institute (2014b) *ICIQ structure Short Form*. Available at: http://www.iciq.net/ICIQ-UIshortform.html (Accessed: 26 June 2018).

Bristol Urological Institute (2014c) *User policy*. Available at: http://www.iciq.net/userpolicy.html (Accessed: 26 June 2018).

Brodribb, W., Zadoroznyj, M., and Dane, A. (2013) 'The views of mothers and GPs about postpartum care in Australian general practice', *BMC Family Practice*, 14, p. 139. doi: 10.1186/1471-2296-14-139.

Brown and Lumley, J. (1998) 'Maternal health after childbirth: results of an Australian population based survey', *BJOG: An International Journal of Obstetrics & Gynaecology*, 105(2), pp. 156-161.

Browne, R.H. (1995) 'On the use of a pilot sample for sample size determination', *Statistics in Medicine*, 14(17), pp. 1933-1940.

Brubaker, L., Richter, H.E., Barber, M.D., Hsu, Y., Rahn, D.D., Menefee, S., Visco, A., Spino, C., Martin, S., and Meikle, S.F. (2013) 'Pelvic floor disorders clinical trials: participant recruitment and retention', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 24(1), pp. 73-79.

Brueton, V.C., Tierney, J., Stenning, S., Harding, S., Meredith, S., Nazareth, I., and Rait, G. (2013) 'Strategies to improve retention in randomised trials', *Cochrane Database of Systematic Reviews*, (12). doi: 10.1002/14651858.MR000032.pub2.

Buchheit, H. (1985) *Die vaginale Akupunktur: phylogenetische und embryogenetische Grundlagen; neurophysiologische und kulturhistorische Aspekte*. Heidelberg: Haug.

Buchner, A., Erdfelder, E., Faul, F., and Lang, A.-G. (2013) *G*Power for Windows* (Version 3.1) [Computer program]. Available at: http://www.gpower.hhu.de/ (Downloaded: 2 April 2018).

Bugge, C., Williams, B., Hagen, S., Logan, J., Glazener, C., Pringle, S., and Sinclair, L. (2013) 'A process for Decision-making after Pilot and feasibility Trials (ADePT): development following a feasibility study of a complex intervention for pelvic organ prolapse', *Trials*, 14, p. 353. doi: 10.1186/1745-6215-14-353.

Bump, R.C., Hurt, W.G., Fantl, J.A., and Wyman, J.F. (1991) 'Assessment of Kegel pelvic muscle exercise performance after brief verbal instruction', *American Journal of Obstetrics & Gynecology*, 165(2), pp. 322-329.

Bundeskanzleramt Österreich (2013a) *Angaben der Hebamme - Geburt*. Available at: https://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2013_II_324/COO_2026_100_2_931550.pdf (Accessed: 29 June 2018).

Bundeskanzleramt Österreich (2013b) *Anzeige der Geburt*. Available at: https://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2013_II_324/COO_2026_100_2_931549.pdf (Accessed: 29 June 2018).

Bundeskanzleramt Rechtsinformationssystem (1994) *Bundesgesetz über den Hebammenberuf (Hebammengesetz - HebG)*. Available at: https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10010804&ShowPrintPreview=True (Accessed: 26 June 2018).

Bundeskanzleramt Rechtsinformationssystem (1996) *Bundesgesetz betreffend Medizinprodukte (Medizinproduktegesetz - MPG)*. Available at: https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10011003 (Accessed: 6 August 2018).

Bundesministerium für Gesundheit (2016) Mutter-Kind-Pass. Wien: Bundesministerium.

Bundesministerium für Gesundheit (2017) *Unser Baby kommt: Begleitbroschüre zum Mutter-Kind-Pass.* 35th edn. Wien: Bundesministerium. Available at: https://www.bmgf.gv.at/home/Gesundheit/Gesundheitsfoerderung_Praevention/Eltern_und_Kind/Mutter_Kind_Pass (Accessed: 23 July 2018).

Burns, N. and Grove, S. (2009) *The practice of nursing research: appraisal, synthesis, and generation of evidence*. 6th edn. St. Louis, MO: Saunders.

Burns, N.P.D. and Grove, S.K. (2007) *Understanding nursing research: building an evidence-based practice*. 4th edn. Philadelphia, PA: Saunders.

Burns, R.B. (2000) *Introduction to research methods*. London: SAGE.

Butej, R. (2010) *Smart Pelvic Floor - Die präventive Kräftigung des Beckenbodens mit Smart Balls.* Unpublished Diplomarbeit. Medizinisch-technische Akademie für den physiotherapeutischen Dienst des Landes Kärnten am Landeskrankenhaus Klagenfurt.

Calvert, M., Blazeby, J., Altman, D.G., Revicki, D.A., Moher, D., and Brundage, M.D. (2013) 'Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension', *JAMA*, 309(8), pp. 814-822.

Cammu, H. and Van Nylen, M. (1998) 'Pelvic floor exercises versus vaginal weight cones in genuine stress incontinence', *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 77(1), pp. 89-93.

Cammu, H., Van Nylen, M., Blockeel, C., Kaufman, L., and Amy, J.-J. (2004) 'Who will benefit from pelvic floor muscle training for stress urinary incontinence?', *American Journal of Obstetrics & Gynecology*, 191(4), pp. 1152-1157.

CampusPharma (no date) *Vielle Vagitrim*. Available at: http://www.campuspharma.se/vara-receptfria-produkter/vielle-vagitrim/ (Accessed: 26 June 2018).

CANTIENICA AG (2018) *CANTIENICA® method for bodyform & posture*. Available at: http://www.cantienica-method.com/ (Accessed: 26 June 2018).

Cardozo, L. and Kelleher, C. (1997) *The King's Health Questionnaire*. Available at: https://www.nice.org.uk/guidance/cg171/resources/urinary-incontinence-the-kings-health-questionnaire2 (Accessed: 26 June 2018).

Caroci, A.d.S., Riesco, M.L., Rocha, B.M., Ventura, L.d.J., and Oliveira, S.G. (2014) 'Evaluation of perineal muscle strength in the first trimester of pregnancy', *Revista Latino-Americana de Enfermagem (RLAE)*, 22(6), pp. 893-901.

Castro, R.A., Arruda, R.M., Zanetti, M.R., Santos, P.D., Sartori, M.G., and Girao, M.J. (2008) 'Single-blind, randomized, controlled trial of pelvic floor muscle training, electrical stimulation, vaginal cones, and no active treatment in the management of stress urinary incontinence', *Clinics*, 63(4), pp. 465-472.

Centers for Disease Control and Prevention (2015) *2015 Sexually transmitted diseases treatment guidelines: vulvovaginal candidiasis*. Available at: https://www.cdc.gov/std/tg2015/candidiasis.htm (Accessed: 26 June 2018).

Centre for Evidence-Based Medicine (2017) *Oxford Centre for Evidence-based Medicine – Levels of evidence (March 2009)*. Available at: http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/ (Accessed: 18 September 2017).

Chaliha, C. (2006) 'Chapter 46: Pregnancy and childbirth and the effect on the pelvic floor', in Cardozo, L. & Staskin, D. (eds.) *Textbook of female urology and urogynecology.* 2nd edn. Abingdon: Informa Healthcare, pp. 681-694.

Chalmers, I. (1995) 'What do I want from health research and researchers when I am a patient?', *BMJ*, 310(6990), pp. 1315-1318.

Chan, A.W., Tetzlaff, J.M., Gotzsche, P.C., Altman, D.G., Mann, H., Berlin, J.A., Dickersin, K., Hrobjartsson, A., Schulz, K.F., Parulekar, W.R., Krleza-Jeric, K., Laupacis, A., and Moher, D. (2013) 'SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials', *BMJ*, 346, p. e7586. doi: 10.1136/bmj.e7586.

'Chapter 5: Measurement of pelvic floor muscle function and strength, and pelvic organ prolapse', (2015) in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* Edinburgh: Churchill Livingstone, pp. 43-109.

Charlesworth, G., Burnell, K., Hoe, J., Orrell, M., and Russell, I. (2013) 'Acceptance checklist for clinical effectiveness pilot trials: a systematic approach', *BMC Medical Research Methodology*, 13, p. 1. doi: 10.1186/1471-2288-13-78.

Chiarelli, P. and Cockburn, J. (2002) 'Promoting urinary continence in women after delivery: randomised controlled trial', *BMJ*, 324(7348), pp. 1241-1244.

Chiarelli, P., Murphy, B., and Cockburn, J. (2003) 'Acceptability of a urinary continence promotion programme to women in postpartum', *BJOG: An International Journal of Obstetrics & Gynaecology*, 110(2), pp. 188-196.

Chiarelli, P., Murphy, B., and Cockburn, J. (2004) 'Promoting urinary continence in postpartum women: 12-month follow-up data from a randomised controlled trial', *International Urogynecology Journal*, 15(2), pp. 99-105.

Chiarelli, P.E. and Moore, K.H. (2008) 'Exercise, feedback, and biofeedback', in Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A. & Stanton, S.L. (eds.) *Pelvic floor re-education: principles and practice.* 2nd edn. London: Springer, pp. 184-189.

City, University of London (2018) *Research governance*. Available at: https://www.city.ac.uk/health/research/research-governance (Accessed: 26 June 2018).

Clark-Carter, D. (2003) 'Effect size-the missing piece in the jigsaw', *Psychologist*, 16(12), pp. 636-638.

Clement, C. (2017) '"To have, to hold, from this day forward": understanding current practice regarding retention of trial participants' [Abstract] (from the Proceedings of the 4th international clinical trials methodology conference, Liverpool, 7-10 May 2017), *Trials*, 18(Suppl 1), p. 159.

Cluett, E.R. and Bluff, R. (2006) *Principles and practice of research in midwifery*. 2nd edn. Edinburgh: Baillière Tindall.

Coad, J. and Dunstall, M. (2011) *Anatomy and physiology for midwives*. Edinburgh: Churchill Livingstone.

Cocks, K. and Torgerson, D.J. (2013) 'Sample size calculations for pilot randomized trials: a confidence interval approach', *Journal of Clinical Epidemiology*, 66(2), pp. 197-201.

Consumers united for evidence-based health care (2011). Available at: https://www.youtube.com/watch?v=q5sySOQA0Mo (Accessed: 26 June 2018).

Cook, D.J., Rocker, G., Meade, M., Guyatt, G., Geerts, W., Anderson, D., Skrobik, Y., Hebert, P., Albert, M., and Cooper, J. (2005) 'Prophylaxis of Thromboembolism in Critical Care (PROTECT) Trial: a pilot study', *Journal of Critical Care*, 20(4), pp. 364-372.

Cooper, C.L., Whitehead, A., Pottrill, E., Julious, S.A., and Walters, S.J. (2018) 'Are pilot trials useful for predicting randomisation and attrition rates in definitive studies: A review of publicly funded trials', *Clinical Trials*, 15(2), pp. 189-196. doi: 10.1177/1740774517752113.

Cornes, H., Bartolo, D.C., and Stirrat, G.M. (1991) 'Changes in anal canal sensation after childbirth', *British Journal of Surgery*, 78(1), pp. 74-77.

Cosner, K.R., Dougherty, M.C., and Bishop, K.R. (1991) 'Dynamic characteristics of the circumvaginal muscles during pregnancy and the postpartum', *Journal of Nurse-Midwifery*, 36(4), pp. 221-225.

Costanzo, L.S. (2010) *Physiology*. 4th edn. Philadelphia, PA: Saunders Elsevier.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., and Petticrew, M. (2006) *Developing and evaluating complex interventions: new guidance*. Available at: www.mrc.ac.uk/complexinterventionsguidance (Accessed: 29 June 2018).

Culligan, P.J., Blackwell, L., Murphy, M., Ziegler, C., and Heit, M.H. (2005) 'A randomized, double-blinded, sham-controlled trial of postpartum extracorporeal magnetic innervation to restore pelvic muscle strength in primiparous patients', *American Journal of Obstetrics & Gynecology*, 192(5), pp. 1578-1582.

Culligan, P.J., Scherer, J., Dyer, K., Priestley, J.L., Guingon-White, G., Delvecchio, D., and Vangeli, M. (2010) 'A randomized clinical trial comparing pelvic floor muscle training to a Pilates exercise program for improving pelvic muscle strength', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 21(4), pp. 401-408.

Cumming, G. (2013) 'The new statistics: why and how', *Psychological Science*, 25(1), pp. 7-29.

Cuzick, J., Edwards, R., and Segnan, N. (1997) 'Adjusting for non-compliance and contamination in randomized clinical trials', *Statistics in Medicine*, 16(9), pp. 1017-1029.

Daykin, A., Clement, C., Gamble, C., Kearney, A., Blazeby, J., Clarke, M., Lane, J.A., and Shaw, A. (2018) ''Recruitment, recruitment, recruitment'–the need for more focus on retention: a qualitative study of five trials', *Trials*, 19(1), p. 76. doi: 10.1186/s13063-018-2467-0.

Deegan, E.G., Stothers, L., Kavanagh, A., and Macnab, A.J. (2018) 'Quantification of pelvic floor muscle strength in female urinary incontinence: a systematic review and comparison of contemporary methodologies', *Neurourology and Urodynamics*, 37(1), pp. 33-45.

Deffieux, X., Vieillefosse, S., Billecocq, S., Battut, A., Nizard, J., Coulm, B., and Thubert, T. (2015) 'Postpartum pelvic floor muscle training and abdominal rehabilitation: guidelines', *Journal de Gynecologie, Obstetrique et Biologie de la Reproduction*, 44(10), pp. 1141-1146.

'Definition of 'clinical'' (2018), in *Collins English dictionary*. Available at: https://www.collinsdictionary.com/dictionary/english/clinical (Accessed: 25 September 2018).

'Definition von cone' (2019), in *Collins*. Available at: https://www.collinsdictionary.com/de/worterbuch/englisch/cone (Accessed: 30 November 2019).

Deindl, F., Schüssler, B., Vodusek, D., and Hesse, U. (1995) 'Neurophysiologic effect of vaginal cone application in continent and urinary stress incontinent women', *International Urogynecology Journal*, 6(4), pp. 204-208.

DeLancey, J.O. (1990) 'Anatomy and physiology of urinary continence', *Clinical Obstetrics and Gynecology*, 33(2), pp. 298-307.

Delancey, J.O. (1993) 'Anatomy and biomechanics of genital prolapse', *Clinical Obstetrics and Gynecology*, 36(4), pp. 897-909.

DeLancey, J.O. (2005) 'The hidden epidemic of pelvic floor dysfunction: achievable goals for improved prevention and treatment', *American Journal of Obstetrics & Gynecology*, 192(5), pp. 1488-1495.

DeLancey, J.O., Kearney, R., Chou, Q., Speights, S., and Binno, S. (2003) 'The appearance of levator ani muscle abnormalities in magnetic resonance images after vaginal delivery', *Obstetrics & Gynecology*, 101(1), pp. 46-53.

DeLancey, J.O., Morgan, D.M., Fenner, D.E., Kearney, R., Guire, K., Miller, J.M., Hussain, H., Umek, W., Hsu, Y., and Ashton-Miller, J.A. (2007) 'Comparison of levator ani muscle defects and function in women with and without pelvic organ prolapse', *Obstetrics & Gynecology*, 109(2 Pt 1), pp. 295-302.

Delancey, J.O.L. (2016) 'True pelvis, pelvic floor and perineum', in Standring, S. & Gray, H. (eds.) *Gray's anatomy: the anatomical basis of clinical practice.* 41st edn. Edinburgh: Churchill Livingstone, pp. 1221-1236.

DeLancey, J.O.L. and Ashton-Miller, J.A. (2015) 'MRI of intact and injured female pelvic floor muscles', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* 2nd edn. Edinburgh: Churchill Livingstone, pp. 97-109.

Demott, K., Bick, D., Norman, R., Ritchie, G., Turnbull, N., Adams, C., Barry, C., Byrom, S., Elliman, D., Marchant, S., Mccandlish, R., Mellows, H., Neale, C., Parkar, M., Tait, P., and Taylor, C. (2006) *Clinical guidelines and evidence review for post natal care: routine postnatal care of recently delivered women and their babies*. Available at: https://www.nice.org.uk/guidance/cg37/evidence/full-guideline-pdf-485782237 (Accessed: 29 June 2018).

Denzin, N. (2009) *The research act: a theoretical introduction to sociological methods*. New Brunswick, NJ: Aldine Transaction.

Deshpande, P.R., Rajan, S., Sudeepthi, B.L., and Abdul Nazir, C.P. (2011) 'Patient-reported outcomes: a new era in clinical research', *Perspectives in Clinical Research*, 2(4), pp. 137-144.

Devreese, A., Staes, F., De Weerdt, W., Feys, H., Van Assche, A., Penninckx, F., and Vereecken, R. (2004) 'Clinical evaluation of pelvic floor muscle function in continent and incontinent women', *Neurourology and Urodynamics*, 23(3), pp. 190-197.

Dias, L.A.R., Driusso, P., Aita, D.L.C.C., Quintana, S.M., Bø, K., and Ferreira, C.H.J. (2011) 'Effect of pelvic floor muscle training on labour and newborn outcomes: a randomized controlled trial', *Revista Brasileira de Fisioterapia*, 15(6), pp. 487-493.

Dietz, H. and Shek, K. (2008a) 'The quantification of levator muscle resting tone by digital assessment', *International Urogynecology Journal*, 19(11), pp. 1489-1493.

Dietz, H.P. (2013) 'Pelvic floor trauma in childbirth', *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 53(3), pp. 220-230.

Dietz, H.P., Clarke, B., and Vancaillie, T.G. (2002) 'Vaginal childbirth and bladder neck mobility', *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 42(5), pp. 522-525.

Dietz, H.P., Kirby, A., Shek, K.L., and Bedwell, P.J. (2009) 'Does avulsion of the puborectalis muscle affect bladder function?', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 20(8), pp. 967-972.

Dietz, H.P. and Lanzarone, V. (2005) 'Levator trauma after vaginal delivery', *Obstetrics & Gynecology*, 106(4), pp. 707-712.

Dietz, H.P. and Schierlitz, L. (2005) 'Pelvic floor trauma in childbirth - myth or reality?', *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 45(1), pp. 3-11.

Dietz, H.P. and Shek, C. (2008b) 'Levator avulsion and grading of pelvic floor muscle strength', *International Urogynecology Journal*, 19(5), pp. 633-636.

Dietz, H.P., Shek, K.L., Chantarasorn, V., and Langer, S.E.M. (2012) 'Do women notice the effect of childbirth-related pelvic floor trauma?', *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 52(3), pp. 277-281.

Dietz, H.P. and Simpson, J.M. (2008) 'Levator trauma is associated with pelvic organ prolapse', *BJOG: An International Journal of Obstetrics & Gynaecology*, 115(8), pp. 979-984.

Dietz, H.P. and Steensma, A.B. (2006) 'The role of childbirth in the aetiology of rectocele', *BJOG: An International Journal of Obstetrics & Gynaecology*, 113(3), pp. 264-267.

Diez-Itza, I., Arrue, M., Ibanez, L., Paredes, J., Murgiondo, A., and Sarasqueta, C. (2011a) 'Influence of mode of delivery on pelvic organ support 6 months postpartum', *Gynecologic and Obstetric Investigation*, 72(2), pp. 123-129.

Diez-Itza, I., Arrue, M., Ibanez, L., Paredes, J., Murgiondo, A., and Sarasqueta, C. (2011b) 'Postpartum impairment of pelvic floor muscle function: factors involved and association with prolapse', *International Urogynecology Journal*, 22(12), pp. 1505-1511.

DiNubile, N. (1991) 'Strength training', *Clinics in Sports Medicine*, 10(1), pp. 33-62.

DocCheck Medical Services GmbH (2018a) *Barorezeptor*. Available at: http://flexikon.doccheck.com/de/Barorezeptor (Accessed: 29 June 2018).

DocCheck Medical Services GmbH (2018b) *Rehabilitation*. Available at: http://flexikon.doccheck.com/de/Rehabilitation (Accessed: 26 June 2018).

DocCheck Medical Services GmbH (2018c) *Vater-Pacini-Körperchen*. Available at: http://flexikon.doccheck.com/de/Vater-Pacini-K%C3%B6rperchen (Accessed: 9 August 2018).

Donovan, J.L., Rooshenas, L., Jepson, M., Elliott, D., Wade, J., Avery, K., Mills, N., Wilson, C., Paramasivan, S., and Blazeby, J.M. (2016) 'Optimising recruitment and informed consent in randomised controlled trials: the development and implementation of the Quintet Recruitment Intervention (QRI)', *Trials*, 17(1), p. 1. doi: 10.1186/s13063-016-1391-4.

Dougherty, M.C., Bishop, K.R., Abrams, R.M., Batich, C.D., and Gimotty, P.A. (1989) 'The effect of exercise on the circumvaginal muscles in postpartum women', *Journal of Nurse-Midwifery*, 34(1), pp. 8-14.

Dougherty, M.C., Bishop, K.R., Mooney, R.A., Gimotty, P.A., and Landy, L.B. (1991) 'Variation in intravaginal pressure measurements', *Nursing Research*, 40(5), pp. 282-285.

Duan, N. (2013) 'From pilot studies to confirmatory studies', *Shanghai Archives of Psychiatry*, 25(5), pp. 325-328.

Duchesnay (2017) *What is LadySystem²* Available at: https://www.ladysystem.ca/en/product (Accessed: 29 June 2018).

Dumoulin, C. (2004) *Efficacité des traitements physiotherapiques pour l'incontinence urinaire d'effort chez la femme en periode postnatale.* PhD thesis. Université de Montréal. Available at: https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/14491/Dumoulin_Chantale_2003_these.pdf?sequence=1&isAllowed=y (Accessed: 29 June 2018).

Dumoulin, C., Adewuyi, T., Booth, J., Bradley, C., Burgio, K., Hagen, S., Hunter, K., Imamura, M., Morin, M., Mørkved, S., Thakar, R., Wallace, S., and Williams, K. (2017) 'Adult conservative management', in Abrams, P., Cardozo, L., Wagg, A. & Wein, A. (eds.) *Incontinence.* 6th edn. No place: ICUD ICS, pp. 1443-1628.

Dumoulin, C., Hay-Smith, E.J.C., and Mac Habée-Séguin, G. (2014) 'Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women', *Cochrane Database of Systematic Reviews*, (5). doi: 10.1002/14651858.CD005654.pub3.

Dumoulin, C., Hunter, K.F., Moore, K., Bradley, C.S., Burgio, K.L., Hagen, S., Imamura, M., Thakar, R., Williams, K., and Chambers, T. (2016) 'Conservative management for female urinary incontinence and pelvic organ prolapse review 2013: summary of the 5th International Consultation on Incontinence', *Neurourology and Urodynamics*, 35(1), pp. 15-20.

Dumoulin, C., Peng, Q., Stodkilde-Jorgensen, H., Shishido, K., and Constantinou, C. (2007) 'Changes in levator ani anatomical configuration following physiotherapy in women with stress urinary incontinence', *The Journal of Urology*, 178(3), pp. 970-977.

Dunkley, J. (2000) *Health promotion in midwifery*. Edinburgh: Baillière Tindall.

Durlak, J.A. (2009) 'How to select, calculate, and interpret effect sizes', *Journal of Pediatric Psychology*, 34(9), pp. 917-928.

Dykes, M. (2003) *Crash course: anatomy*. 2nd edn. Edinburgh: Mosby.

Edwards, P.J., Roberts, I., Clarke, M.J., DiGuiseppi, C., Wentz, R., Kwan, I., Cooper, R., Felix, L.M., and Pratap, S. (2009) 'Methods to increase response to postal and electronic questionnaires', *Cochrane Database of Systematic Reviews*, (3). doi: 10.1002/14651858.MR000008.pub4.

Efird, J. (2011) 'Blocked randomization with randomly selected block sizes', *International Journal of Environmental Research and Public Health*, 8(1), pp. 15-20.

Efron, B. and Feldman, D. (1991) 'Compliance as an explanatory variable in clinical trials', *Journal of the American Statistical Association*, 86(413), pp. 9-17.

ELANEE (2017a) E-mail to Claudia Oblasser. 24 August.

ELANEE (2017b) *Pelvic floor exercise aid phase II*. Available at: http://www.elanee.de/en/startseite/beckenbodentraining/beckenbodentrainingshilfe-phase-ii/ (Accessed: 26 June 2018).

ELANEE (2017c) *Pelvic floor exercise aids phase I active*. Available at: http://www.elanee.de/en/startseite/beckenbodentraining/beckenbodentrainingshilfe-phase-i-aktiv/ (Accessed: 26 June 2018).

Eldridge, S. (2010) 'Pragmatic trials in primary health care: what, when and how?', *Family Practice*, 27(6), pp. 591-592.

Eldridge, S. (2013) 'What is a pilot study and how do I do a good one?' [PowerPoint Presentation]. London.

Eldridge, S. and Kerry, S. (2012) *A practical guide to cluster randomised trials in health services research*. Chichester: Wiley.

Eldridge, S.M., Chan, C.L., Campbell, M.J., Bond, C.M., Hopewell, S., Thabane, L., and Lancaster, G.A. (2016a) 'CONSORT 2010 statement: extension to randomised pilot and feasibility trials', *BMJ*, 355, p. i5239. doi: 10.1136/bmj.i5239.

Eldridge, S.M., Lancaster, G.A., Campbell, M.J., Thabane, L., Hopewell, S., Coleman, C.L., and Bond, C.M. (2016b) 'Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework', *PLoS One*, 11(3), p. e0150205. doi: 10.1371/journal.pone.0150205 (Accessed: 11 August 2018).

Encyclopædia Britannica, Inc. (2017) *Acrylonitrile-butadiene-styrene copolymer (ABS)*. Available at: https://www.britannica.com/science/acrylonitrile-butadiene-styrene-copolymer (Accessed: 4 July 2018).

Entwistle, V.A., Renfrew, M.J., Yearley, S., Forrester, J., and Lamont, T. (1998) 'Lay perspectives: advantages for health research', *BMJ*, 316(7129), pp. 463-466.

Erol, B., Sanli, O., Korkmaz, D., Seyhan, A., Akman, T., and Kadioglu, A. (2007) 'A cross-sectional study of female sexual function and dysfunction during pregnancy', *The Journal of Sexual Medicine*, 4(5), pp. 1381-1387.

Esberger, M. (2007) *Säuglingsernährung heute 2006: Struktur- und Beratungsqualität an den Geburtenkliniken in Österreich Ernährung von Säuglingen im ersten Lebensjahr*. Available at: http://www.richtigessenvonanfangan.at/fileadmin/Redakteure_REVAN/user_upload/Langversion_Studie_zum_Thema_Stillen_und_dem_Ern%C3%A4hrungsverhalten_von_S%C3%A4uglingen.pdf (Accessed: 29 June 2018).

European Commission (2018) *CE marking*. Available at: https://ec.europa.eu/growth/single-market/ce-marking_en (Accessed: 6 August 2018).

European Medicines Agency (1995) *Note for guidance on clinical safety data management: definitions and standards for expedited reporting (CPMP/ICH/377/95)*. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002749.pdf (Accessed: 29 June 2018).

Ewings, P., Spencer, S., Marsh, H., and O'Sullivan, M. (2005) 'Obstetric risk factors for urinary incontinence and preventative pelvic floor exercises: cohort study and nested randomized controlled trial', *Journal of Obstetrics and Gynaecology*, 25(6), pp. 558-564.

Farlex, Inc. (2018a) *Antepartum*. Available at: https://medical-dictionary.thefreedictionary.com/antepartum (Accessed: 13 September 2018).

Farlex, Inc. (2018b) *Biofeedback*. Available at: http://medical-dictionary.thefreedictionary.com/biofeedback (Accessed: 26 June 2018).

Farlex, Inc. (2018c) *Concentric*. Available at: https://medical-dictionary.thefreedictionary.com/concentric (Accessed: 29 June 2018).

Farlex, Inc. (2018d) *Eccentric*. Available at: https://medical-dictionary.thefreedictionary.com/eccentric (Accessed: 29 June 2018).

Farlex, Inc. (2018e) *Electromyography*. Available at: https://medical-dictionary.thefreedictionary.com/electromyography (Accessed: 22 August 2018).

Farlex, Inc. (2018f) *Ischaemia*. Available at: http://medical-dictionary.thefreedictionary.com/ischaemia (Accessed: 26 June 2018).

Farlex, Inc. (2018g) *Isometric*. Available at: http://medical-dictionary.thefreedictionary.com/isometric (Accessed: 26 June 2018).

Farlex, Inc. (2018h) *Nocturia*. Available at: https://medical-dictionary.thefreedictionary.com/nocturia (Accessed: 9 July 2018).

Farlex, Inc. (2018i) *Pacinian corpuscle*. Available at: https://www.thefreedictionary.com/Vater-pacini+corpuscle (Accessed: 9 August 2018).

Farlex, Inc. (2018j) *Postural*. Available at: http://medical-dictionary.thefreedictionary.com/postural (Accessed: 26 June 2018).

Farlex, Inc. (2018k) *Proprioceptor*. Available at: https://medical-dictionary.thefreedictionary.com/proprioceptor (Accessed: 31 July 2018).

Farlex, Inc. (2018l) *Reflexive*. Available at: http://www.thefreedictionary.com/Reflexive (Accessed: 26 June 2018).

Farlex, Inc. (2018m) *Secondary prevention*. Available at: http://medical-dictionary.thefreedictionary.com/secondary+prevention (Accessed: 26 June 2018).

Farlex, Inc. (2018n) *Postnatal*. Available at: http://medical-dictionary.thefreedictionary.com/postnatal (Accessed: 26 June 2018).

Federal Communications Commission (no date) *About the FCC*. Available at: https://www.fcc.gov/about/overview (Accessed: 29 June 2018).

Federative International Programme on Anatomical Terminologies (2011) *Terminologia anatomica: international anatomical terminology*. 2nd edn. Stuttgart: Georg Thieme.

'Femcon-Vaginalkonen' (2013), *Österreichische Hebammenzeitung*, 19(6), p. 25.

Fergusson, D., Aaron, S.D., Guyatt, G., and Hébert, P. (2002) 'Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis', *BMJ*, 325(7365), pp. 652-654.

Fidler, F. (2002) 'The fifth edition of the APA Publication Manual: why its statistics recommendations are so controversial', *Educational and Psychological Measurement*, 62(5), pp. 749-770.

Fine, P., Burgio, K., Borello-France, D., Richter, H., Whitehead, W., Weber, A., and Brown, M. (2007) 'Teaching and practicing of pelvic floor muscle exercises in primiparous women during pregnancy and the postpartum period [corrected; published erratum in American Journal of Obstetrics & Gynecology 2007, 197(3), p. 333]', *American Journal of Obstetrics & Gynecology*, 197(1), pp. 107.e1-107.e5.

Finlay, H. (2006) 'Public health, midwifery and government policy', in Bowden, J. & Manning, V. (eds.) *Health promotion in midwifery: principles and practice.* London: Hodder Arnold, pp. 1-12.

Fischer, W. and Baessler, K. (1996) 'Postpartum pelvic floor conditioning using vaginal cones: not only for prophylaxis against urinary incontinence and descensus', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 7(4), pp. 208-214.

Fischer, W., Baessler, K., and Linde, A. (1996) 'Beckenbodenkonditionierung mit Vaginalgewichten – post partum und bei Harninkontinenz', *Zentralblatt für Gynäkologie*, 118(1), pp. 18-28.

Fishpond Ltd (2018) *2.5cm golden Kegel Ben Wa Balls with passion lubricant*. Available at: https://www.fishpond.co.nz/Health/25cm-Golden-Kegel-Ben-Wa-Balls-with-Passion-Lubricant/0848518017246 (Accessed: 26 June 2018).

Fleck, S.J. and Falkel, J.E. (1986) 'Value of resistance training for the reduction of sports injuries', *Sports Medicine*, 3(1), pp. 61-68.

FOKUS KIND Medien (2017) *Nachsorge*. Available at: https://www.schwanger.at/artikel/nachsorge.html (Accessed: 26 June 2018).

Foldspang, A., Mommsen, S., and Djurhuus, J.C. (1999) 'Prevalent urinary incontinence as a correlate of pregnancy, vaginal childbirth, and obstetric techniques', *American Journal of Public Health*, 89(2), pp. 209-212.

Forman, R. (1981) 'Medical resistance to innovation', *Medical Hypotheses*, 7(8), pp. 1009-1017.

Foxman, B., Muraglia, R., Dietz, J.-P., Sobel, J.D., and Wagner, J. (2013) 'Prevalence of recurrent vulvovaginal candidiasis in 5 european countries and the united states: results from an internet panel survey', *Journal of Lower Genital Tract Disease*, 17(3), pp. 340-345.

Franke, T. (2010) 'Wochenbett- und Rückbildungsgymnastik ist mehr als Prävention gegen eine spätere Inkontinenz', *Hebammenforum*, (1). Available at: http://docplayer.org/11639079-Wochenbett-und-rueckbildungsgymnastik-ist-mehr-als-praevention-gegen-eine-spaetere-inkontinenz.html (Accessed: 26 June 2018).

Frawley, H. (2006) 'Pelvic floor muscle strength testing', *Australian Journal of Physiotherapy*, 52(4), p. 307.

Frawley, H., Galea, M., Phillips, B., Sherburn, M., and Bo, K. (2004) 'Effect of different testing positions using three measurement tools on pelvic floor muscle contraction, and subject acceptance of testing position' [Abstract], *International Continence Society/International Urogynecological Association annual meeting*, Paris, 25-27 August. Available at: https://www.ics.org/2004/programme?q=Effect%20of%20different%20testing%20positions%20using%20three%20measurement%20tools%20on%20pelvic%20floor%20muscle%20contraction,%20and%20subject%20acceptance%20of%20testing%20position (Accessed: 26 June 2018).

Frawley, H.C., Galea, M.P., Phillips, B.A., Sherburn, M., and Bø, K. (2006) 'Reliability of pelvic floor muscle strength assessment using different test positions and tools', *Neurourology and Urodynamics*, 25(3), pp. 236-242.

Friedman, S., Blomquist, J.L., Nugent, J.M., McDermott, K.C., Munoz, A., and Handa, V.L. (2012) 'Pelvic muscle strength after childbirth', *Obstetrics & Gynecology*, 120(5), pp. 1021-1028.

Fritsch, H. (2012) 'Anatomie und Physiologie des Beckenbodens', in Carrière, B. (ed.) *Beckenboden: Physiotherapie und Training.* 2nd edn. Stuttgart: Thieme, pp. 1-20.

Fry, C., Chess-Williams, R., Hashitani, H., Kanai, A., McCloskey, K., Takeda, M., and Vahabi, B. (2017) 'Cell biology', in Abrams, P., Cardozo, L., Wagg, A. & Wein, A. (eds.) *Incontinence.* 6th edn. No place: ICUD ICS, pp. 143-258.

FUN FACTORY (2013) *User manual.* Available at: https://www.funfactory.com/fileadmin/manuals/fun_factory_um_smartballs_uno_duo.pdf (Accessed: 3 August 2018).

FUN FACTORY (2014) E-mail to Claudia Oblasser. 24 November.

FUN FACTORY (2018) E-mail to Claudia Oblasser. 8 January.

FUN FACTORY (no date-a) *55 facts of Smartballs.* Available at: https://www.funfactory.com/fileadmin/manuals/Booklet_55Facts_of_SMARTBALLS_deutsch_view.pdf (Accessed: 3 August 2018).

FUN FACTORY (no date-b) *Smartball duo*. Available at: https://www.funfactory.com/de/liebeskugeln/smartballs-duo/ (Accessed: 3 August 2018).

FUN FACTORY (no date-c) *Smartball duo*. Available at: https://www.funfactory.com/en/smartballs/smartballs-duo/ (Accessed: 3 August 2018).

FUN FACTORY (no date-d) *Smartball uno*. Available at: https://www.funfactory.com/en/smartballs/smartball-uno/ (Accessed: 3 August 2018).

FUN FACTORY (no date-e) *Smartball uno*. Available at: https://www.funfactory.com/de/liebeskugeln/smartball-uno/ (Accessed: 3 August 2018).

FUN FACTORY (no date-f) *Smartballs Produkt Information.* [Leaflet obtained from FUN FACTORY], 6 July 2018.

Gamble, C., Krishan, A., Stocken, D., Lewis, S., Juszczak, E., Doré, C., Williamson, P.R., Altman, D.G., Montgomery, A., and Lim, P. (2017) 'Guidelines for the content of statistical analysis plans in clinical trials', *JAMA*, 318(23), pp. 2337-2343.

Gameiro, M.O., Moreira, E.H., Gameiro, F.O., Moreno, J.C., Padovani, C.R., and Amaro, J.L. (2010) 'Vaginal weight cone versus assisted pelvic floor muscle training in the treatment of female urinary incontinence. A prospective, single-blind, randomized trial', *International Urogynecology Journal*, 21(4), pp. 395-399.

Gardner, E.P. and Johnson, K.O. (2013) 'The somatosensory system: receptors and central pathways', in Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S.A. & Hudspeth, A.J. (eds.) *Principles of neural science.* 5th edn. New York, NY: McGraw-Hill Companies, Inc., pp. 475-497.

Gätje, R., Eberle, C., Scholz, C., Lübke, M., Solbach, C., Muschel, K., Kissler, S., Siedentopf, F., Weißenbacher, T., Debus, G., Schuhmacher, I., and Sänger, N. (2011) *Kurzlehrbuch Gynäkologie und Geburtshilfe*. Stuttgart: Georg Thieme Verlag.

Geburtsallianz Österreich (no date) *Mutter-Kind-Pass*. Available at: http://www.geburtsallianz.at/mutter-kind-pass/ (Accessed: 26 June 2018).

Geist, C. (2005) 'Wochenbettbetreuung in der Klinik', in Geist, C., Harder, U. & Stiefel, A. (eds.) *Hebammenkunde.* 3rd edn. Stuttgart: Hippokrates, pp. 420-426.

Gerrard, J. and ten Hove, R. (2013) *RCM/CSP joint statement on pelvic floor muscle exercise. Improving health outcomes for women following pregnancy and birth*. Available at: https://www.rcm.org.uk/sites/default/files/CSP-000924_RCM.PDF (Accessed: 29 June 2018).

Gerrish, K. and Lathlean, J. (2015) *Research process in nursing*. Chichester: Wiley.

'Gib dir die Kugel!', (no date) *magnolias*, pp. 18-19.

Gibbs, G. (2012) *Analyzing qualitative data*. Los Angeles, CA: Sage.

Gillard, S. and Shamley, D. (2010) 'Factors motivating women to commence and adhere to pelvic floor muscle exercises following a perineal tear at delivery: the influence of experience', *Journal of the Association of Chartered Physiotherapists in Women's Health*, 106, pp. 5-18.

Glavind, K. (2001) 'Conservative treatment of stress incontinence with Geisha balls', *International Urogynecology Journal*, 12(4), pp. 223-224.

Glavind, K. (2013) E-mail to Claudia Oblasser. 5 March.

Glerean, M., Furci, A., Galich, A.M., Fama, B., and Plantalech, L. (2010) 'Bone and mineral metabolism in primiparous women and its relationship with breastfeeding: a longitudinal study', *Medicina*, 70(3), pp. 227-232.

Gödl-Purrer, B. (2006) 'Training and functional exercises for the muscles of the pelvic floor', in Carrière, B. & Markel Feldt, C. (eds.) *The pelvic floor.* Stuttgart: Thieme, pp. 251-266.

Goetghebeur, E.J. and Shapiro, S.H. (1996) 'Analysing non-compliance in clinical trials: ethical imperative or mission impossible?', *Statistics in Medicine*, 15(24), pp. 2813-2826.

Gonçalves, B., Ferreira, C., Alves, C.T., Henriques, M., Azeredo, J., and Silva, S. (2016) 'Vulvovaginal candidiasis: epidemiology, microbiology and risk factors', *Critical Reviews in Microbiology*, 42(6), pp. 905-927.

Gorbea Chavez, V., Velazquez Sanchez Mdel, P., and Kunhardt Rasch, J.R. (2004) 'Efecto de los ejercicios del piso pelvico durante el embarazo y el puerperio en la prevencion de la incontinencia urinaria de esfuerzo', *Ginecologia y Obstetricia de Mexico*, 72(12), pp. 628-636.

Graber, B. and Kline-Graber, G. (1979) 'Female orgasm: role of pubococcygeus muscle', *Journal of Clinical Psychiatry*, 40(8), pp. 348-351.

Gravel, J., Opatrny, L., and Shapiro, S. (2007) 'The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say?', *Clinical Trials*, 4(4), pp. 350-356.

Griffin, C., Dougherty, M.C., and Yarandi, H. (1994) 'Pelvic muscles during rest: responses to pelvic muscle exercise', *Nursing Research*, 43(3), pp. 164-167.

Grissom, R.J. and Kim, J.J. (2012) *Effect sizes for research: univariate and multivariate applications*. 2nd edn. New York, NY: Routledge.

Guerrero, K., Owen, L., Hirst, G., and Emery, S. (2007) 'Antenatal pelvic floor exercises: a survey of both patients' and health professionals' beliefs and practice', *Journal of Obstetrics and Gynaecology*, 27(7), pp. 684-687.

Guise, J.M., Morris, C., Osterweil, P., Li, H., Rosenberg, D., and Greenlick, M. (2007) 'Incidence of fecal incontinence after childbirth', *Obstetrics & Gynecology*, 109(2 Pt 1), pp. 281-288.

Gupta, S.K. (2011) 'Intention-to-treat concept: a review', *Perspectives in Clinical Research*, 2(3), pp. 109-112.

Guyatt, G.H., Mills, E.J., and Elbourne, D. (2008) 'In the era of systematic reviews, does the size of an individual trial still matter', *PLoS Medicine*, 5(1), p. e4. doi: 10.1371/journal.pmed.0050004.

Hägglund, D. and Ahlström, G. (2007) 'The meaning of women's experience of living with long-term urinary incontinence is powerlessness', *Journal of Clinical Nursing*, 16(10), pp. 1946-1954.

Hahn, I., Milsom, I., Ohlsson, B.L., Ekelund, P., Uhlemann, C., and Fall, M. (1996) 'Comparative assessment of pelvic floor function using vaginal cones, vaginal digital palpation and vaginal pressure measurements', *Gynecologic and Obstetric Investigation*, 41(4), pp. 269-274.

Halpern, S.D.M., Karlawish, J.H.T.M.D., and Berlin, J.A.S. (2002) 'The continuing unethical conduct of underpowered clinical trials', *JAMA*, 288(3), pp. 358-362.

Handa, V.L., Cundiff, G., Chang, H.H., and Helzlsouer, K.J. (2008) 'Female sexual function and pelvic floor disorders', *Obstetrics & Gynecology*, 111(5), pp. 1045-1052.

Handa, V.L., Nygaard, I., Kenton, K., Cundiff, G.W., Ghetti, C., Ye, W., and Richter, H.E. (2009) 'Pelvic organ support among primiparous women in the first year after childbirth', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 20(12), pp. 1407-1411.

Handa, V.L., Zyczynski, H.M., Burgio, K.L., Fitzgerald, M.P., Borello-France, D., Janz, N.K., Fine, P.M., Whitehead, W., Brown, M.B., and Weber, A.M. (2007) 'The impact of fecal and urinary incontinence on quality of life 6 months after childbirth', *American Journal of Obstetrics & Gynecology*, 197(6), pp. 636.e1-636.e6.

Hanley, J.A. and Lippman-Hand, A. (1983) 'If nothing goes wrong, is everything all right? Interpreting zero numerators', *JAMA*, 249(13), pp. 1743-17455.

Hanzal, E., Bartosch, B., Stelzhammer, C., and Udier, E. (2015) *Palpation für das Beckenbodentraining*. Berlin: Walter de Gruyter.

Harder, U., Friedrich, J., Kirchner, S., Polleit, H., and Stiefel, A. (2014) *Wochenbettbetreuung in der Klinik und zu Hause*. Stuttgart: Georg Thieme.

Harpe, S.E. (2015) 'How to analyze Likert and other rating scale data', *Currents in Pharmacy Teaching and Learning*, 7(6), pp. 836-850.

Harvey, M.A. (2003) 'Pelvic floor exercises during and after pregnancy: a systematic review of their role in preventing pelvic floor dysfunction', *Journal of Obstetrics and Gynaecology Canada*, 25(6), pp. 487-498.

Haskell, W.L. (2012) 'Dose-response ssues in physical activity, fitness and health', in Bouchard, C., Blair, S.N. & Haskell, W.L. (eds.) *Physical activity and health.* 2nd edn. Champaign, IL: Human Kinetics, pp. 345-358.

Hay-Smith, E.J.C., Herderschee, R., Dumoulin, C., and Herbison, G.P. (2011) 'Comparisons of approaches to pelvic floor muscle training for urinary incontinence in women', *Cochrane Database of Systematic Reviews*, (12). doi: 10.1002/14651858.CD009508.

Haylen, B.T., de Ridder, D., Freeman, R.M., and Swift, S.E. (2010) 'An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction', *Neurourology and Urodynamics*, 29(1), pp. 4-20.

Hays, R. and Hadorn, D. (1992) 'Responsiveness to change: an aspect of validity, not a separate dimension', *Quality of Life Research*, 1(1), pp. 73-75.

Health Research Council of New Zealand (no date) *Feasibility studies*. Available at: http://www.hrc.govt.nz/funding-opportunities/researcher-initiated-proposals/feasibility-studies (Accessed: 23 July 2018).

Hebamme4U (no date) *Beckenboden und Liebeskugeln*. Available at: https://www.hebamme4u.net/forum/baby-ist-da/beckenboden-und-liebeskugeln-t29022.html (Accessed: 23 September 2018).

Heller, A. (2015) *Nach der Geburt: Wochenbett und Rückbildung*. 2nd edn. Stuttgart: Georg Thieme.

Heller, A. (no date) *Bewegungsübungen zur Rückbildung nach der Geburt.* [Leaflet obtained in Vienna General Hospital], August 2014.

Herbison, P.G. and Dean, N. (2013) 'Weighted vaginal cones for urinary incontinence', *Cochrane Database of Systematic Reviews*, (7). doi: 10.1002/14651858.CD002114.pub2.

Herman & Wallace (2018) *Vagina balls*. Available at: https://hermanwallace.com/blog/vagina-balls (Accessed: 26 June 2018).

Hertzog, M.A. (2008) 'Considerations in determining sample size for pilot studies', *Research in Nursing & Health*, 31(2), pp. 180-191.

Higgins, J.P.T. and Altman, D.G. (2008) 'Chapter 8: Assessing risk of bias in included studies', in Higgins, J.P.T. & Green, S. (eds.) *Cochrane handbook for systematic reviews of interventions.* Chichester: Wiley-Blackwell, pp. 187-241.

Higgins, J.P.T., Altman, D.G., and Sterne, J.A.C. (2011a) 'Chapter 8: Assessing risk of bias in included studies', in Higgins, J.P.T. & Green, S. (eds.) *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*. No place: The Cochrane Collaboration. Available at: http://handbook-5-1.cochrane.org/ (Accessed: 29 July 2018).

Higgins, J.P.T., Deeks, J.J., and Altman, D.G. (2011b) 'Chapter 16: Special topics in statistics', in Higgins, J.P.T. & Green, S. (eds.) *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*. No place: The Cochrane Collaboration. Available at: http://handbook-5-1.cochrane.org/ (Accessed: 29 June 2018).

Higgins, J.P.T. and Green, S. (eds.) (2011) *Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated March 2011)*. No place: The Cochrane Collaboration. Available at: http://handbook-5-1.cochrane.org/ (Accessed: 29 June 2018).

Hilde, G., Stær-Jensen, J., Engh, M.E., Brækken, I.H., and Bø, K. (2012) 'Continence and pelvic floor status in nulliparous women at midterm pregnancy', *International Urogynecology Journal*, 23(9), pp. 1257-1263.

Hilde, G., Stær-Jensen, J., Siafarikas, F., Ellström Engh, M., and Bø, K. (2013a) 'Postpartum pelvic floor muscle training and urinary incontinence: a randomized controlled trial', *Obstetrics & Gynecology*, 122(6), pp. 1231-1238.

Hilde, G., Staer-Jensen, J., Siafarikas, F., Engh, M.E., Braekken, I.H., and Bø, K. (2013b) 'Impact of childbirth and mode of delivery on vaginal resting pressure and on pelvic floor muscle strength and endurance', *American Journal of Obstetrics & Gynecology*, 208(1), pp. 50.e1-50.e7. doi: 10.1016/j.ajog.2012.10.878.

Hjartardottir, S., Nilsson, J., Petersen, C., and Lingman, G. (1997) 'The female pelvic floor: a dome-not a basin', *Acta Obstetricia et Gynecologica Scandinavica*, 76(6), pp. 567-571.

Hodges, P., Sapsford, R., and Pengel, L. (2007) 'Postural and respiratory functions of the pelvic floor muscles', *Neurourology and Urodynamics*, 26(3), pp. 362-371.

Hoffmann, T.C., Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S.E., Dixon-Woods, M., McCulloch, P., Wyatt, J., Chan, A.-W., and Michie, S. (2014) 'Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide', *BMJ*, 348, p. g1687. doi: 10.1136/bmj.g1687.

Holland & Barrett Retail Limited (2017) *Lelo Luna Beads*. Available at: http://www.hollandandbarrett.com/shop/product/lelo-luna-beads-60091366 (Accessed: 29 June 2018).

Hollis, S. and Campbell, F. (1999) 'What is meant by intention to treat analysis? Survey of published randomised controlled trials', *BMJ*, 319(7211), pp. 670-674.

Hollmann, W. and Strüder, H.K. (2009) *Sportmedizin: Grundlagen für körperliche Aktivität, Training und Präventivmedizin*. 5th edn. Stuttgart: Schattauer Verlag.

Hooper, R. (no date) *Justifying sample size for a feasibility study*. Available at: http://www.rds-london.nihr.ac.uk/RDSLondon/media/RDSContent/files/PDFs/Justifying-Sample-Size-for-a-Feasibility-Study.pdf (Accessed: 29 June 2018).

Hopewell, S., Clarke, M., Moher, D., Wager, E., Middleton, P., Altman, D.G., Schulz, K.F., and the CONSORT Group (2008) 'CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration', *PLoS Med*, 5(1), p. e20. doi: 10.1371/journal.pmed.0050020.

Horak, T.A., Guzman-Rojas, R.A., Shek, K.L.L., and Dietz, H.P. (2014) 'Pelvic floor trauma: does the second baby matter?', *Ultrasound in Obstetrics & Gynecology*, 44(1), pp. 90-94.

Horne, R., Weinman, J., Barber, N., Elliott, R., Morgan, M., Cribb, A., and Kellar, I. (2005) *Concordance, adherence and compliance in medicine taking*. Available at: http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1412-076_V01.pdf (Accessed: 23 July 2018).

HOT Productions & Vertriebs GmbH (2017) *COME Intimate Care*. Available at: http://www.come-info.at/ (Accessed: 26 June 2018).

how2stats (2014) *Welch's t-test (unequal variances) - SPSS*. Available at: https://www.youtube.com/watch?v=fl6wgE4Edqw (Accessed: 23 June 2018).

how2stats (2015) *What is Cronbach's Alpha? - Explained simply*. Available at: https://www.youtube.com/watch?v=PCztXEfNJLM (Accessed: 23 July 2018).

how2stats (no date) *Confidence intervals for correlations - calculator*. Available at: http://www.how2stats.net/2011/09/confidence-intervals-for-correlations.html (Accessed: 7 September 2018).

Hulley, S.B., Cummings, S.R., Browner, W.S., Grady, D., and Newman, T.B. (2013) *Designing clinical research*. 4th edn. Philadelphia, PA: Lippincott Williams & Wilkins.

Hundley, A.F., Wu, J.M., and Visco, A.G. (2005) 'A comparison of perineometer to brink score for assessment of pelvic floor muscle strength', *American Journal of Obstetrics & Gynecology*, 192(5), pp. 1583-1591.

IBM Corporation (2012/13) *IBM SPSS Statistics for Windows*. Armonk, NY: IBM Corp.

Imamura, M., Abrams, P., Bain, C., Buckley, B., Cardozo, L., Cody, J., Cook, J., Eustice, S., Glazener, C., Grant, A., Hay-Smith, J., Hislop, J., Jenkinson, D., Kilonzo, M., Nabi, G., N'Dow, J., Pickard, R., Ternent, L., Wallace, S., Wardle, J., Zhu, S., and Vale, L. (2010) 'Systematic review and economic modelling of the effectiveness and cost-effectiveness of non-surgical treatments for women with stress urinary incontinence', *Health Technology Assessment*, 14(40). doi: 10.3310/hta14400.

Imperial College London (2018) *PPI resource hub*. Available at: https://www.imperial.ac.uk/patient-experience-research-centre/ppi-resource-hub/ (Accessed: 4 September 20018).

IncoStress (2006) *Tell me doc what exactly is the pelvis?* Available at: https://incostress1.wordpress.com/tag/how-many-bones-in-the-pelvis/ (Accessed: 26 June 2018).

Indrayan, A. (no date) *Varieties of bias to guard against*. Available at: http://www.medicalbiostatistics.com/ (Accessed: 26 June 2018).

Institut für angewandte Erkenntnistheorie und medizinische Methodologie e.V. (2003) *Cognition-based medicine*. Berlin: Springer. Available at: http://www.ifaemm.de/G2_cog.htm.

Institut für klinische Epidemiologie der Tirol Kliniken GmbH (2016) *Geburtenregister Österreich: Bericht über die Geburtshilfe in Österreich 2015*. Available at: https://www.iet.at/page.cfm?vpath=register/geburtenregister/jahresberichte (Accessed: 26 June 2018).

Institut für klinische Epidemiologie der Tirol Kliniken GmbH (2017a) *Geburtenregister Österreich: Bericht über die Geburtshilfe in Österreich 2016*. Available at: https://www.iet.at/page.cfm?vpath=register/geburtenregister/jahresberichte (Accessed: 26 June 2018).

Institut für klinische Epidemiologie der Tirol Kliniken GmbH (2017b) Phone conversation with Claudia Oblasser. 10 November.

Institut für klinische Epidemiologie der Tirol Kliniken GmbH (2018) E-mail to Claudia Oblasser. 14 March.

'Intention to treat analysis and per protocol analysis: complementary information' (2012), *Prescrire International*, 21(133), pp. 304-306.

International Confederation of Midwives (2013) *Essential competencies for basic midwifery practice*. Available at: http://internationalmidwives.org/what-we-do/education-coredocuments/essential-competencies-basic-midwifery-practice/ (Accessed: 23 July 2018).

International Confederation of Midwives (2017) *ICM international definition of the midwife*. The Hague: International Confederation of Midwives. Available at: http://internationalmidwives.org/who-we-are/policy-and-practice/icm-international-definition-of-the-midwife/ (Accessed: 23 July 2018).

International Consortium for Health Outcomes Measurement (2018a) *How we work*. Available at: http://www.ichom.org/how-we-work/ (Accessed: 26 June 2018).

International Consortium for Health Outcomes Measurement (2018b) *ICHOM's mission*. Available at: http://www.ichom.org/why-we-do-it/ (Accessed: 26 June 2018).

International Consortium for Health Outcomes Measurement (2018c) *Pregnancy & childbirth data collection reference guide.* Available at: http://www.ichom.org/medical-conditions/pregnancy-and-childbirth/ (Accessed: 26 June 2018).

Intimina (2015) E-mail to Claudia Oblasser. 15 March.

Intimina (2018a) *Experts in feminine care*. Available at: https://www.intimina.com/en/Who-We-Are.php (Accessed: 6 August 2018).

Intimina (2018b) *Laselle Kegel-Vaginalkugeln.* Available at: https://www.intimina.com/image/mediaMaterial/pdf/Laselle_leaflet_DE.pdf (Accessed: 6 August 2018).

Intimina (2018c) *Laselle training guide (tab How to Use)*. Available at: https://www.intimina.com/en/laselle_weighted_exercisers (Accessed: 6 August 2018).

Intimina (2018d) *Laselle Weighted Exercisers*. Available at: http://www.intimina.com/en/kegel-exercisers-laselle.php (Accessed: 6 August 2018).

Intimina (2018e) *Laselle™ Weighted Exercisers*. Available at: https://www.intimina.com/en/Download-Images.php (Accessed: 6 August 2018).

Intimina (2018f) *What materials are the exercisers made from, are they safe to put inside my body? (tab FAQ)*. Available at: https://www.intimina.com/en/laselle_weighted_exercisers (Accessed: 6 August 2018).

Intimina (2018g) *Why are pelvic floor muscles important? (tab FAQ)*. Available at: https://www.intimina.com/en/laselle_weighted_exercisers (Accessed: 6 August 2018).

Intimina (no date) *Laselle Kegel Exercisers: user manual*. Available at: https://www.intimina.com/image/mediaMaterial/pdf/Manual_de_usuario_de_ejercitadores_de_Kegel_Laselle.pdf (Accessed: 6 August 2018).

INVOLVE (2012) *Briefing notes for researchers: public involvement in NHS, public health and social care research*. Eastleigh: INVOLVE.

Ioannidis, J.P., Evans, S.J., Gotzsche, P.C., O'Neill, R.T., Altman, D.G., Schulz, K., and Moher, D. (2004) 'Better reporting of harms in randomized trials: an extension of the CONSORT statement', *Annals of Internal Medicine*, 141(10), pp. 781-788.

Iosif, S. (1981) 'Stress incontinence during pregnancy and in puerperium', *International Journal of Gynecology & Obstetrics*, 19(1), pp. 13-20.

Jackson, P. (2011) 'Content and organization of postnatal care', in Macdonald, S., Magill-Cuerden, J., Henderson, C. & Mayes, M. (eds.) *Mayes' midwifery.* 14th edn. Edinburgh: Bailliere Tindall, pp. 725-743.

Jacobson, L., Kaij, L., and Nilsson, Å. (1967) 'The course and outcome of the post-partum period from a gynaecological and general somatic standpoint', *Acta Obstetricia et Gynecologica Scandinavica*, 46(2), pp. 183-203.

Jadad, A.R. and Enkin, M. (2007) *Randomized controlled trials: questions, answers, and musings*. 2nd edn. Oxford: Blackwell Publishing Ltd.

Jager, K., Zoccali, C., Macleod, A., and Dekker, F. (2008) 'Confounding: what it is and how to deal with it', *Kidney International*, 73(3), pp. 256-260.

Johnson, S.R., Griffiths, H., and Humberstone, F.J. (2010) 'Attitudes and experience of women to common vaginal infections', *Journal of Lower Genital Tract Disease*, 14(4), pp. 287-294.

Jonasson, A., Fianu, S., and Larsson, B. (1992) 'Vagitrim - nytt, enkelt jhälpmedel för effektiv träning av bäckenbottenmuskulatur', *Lakartidningen*, 89(36), p. 2820.

Jonasson, A., Larsson, B., and Pschera, H. (1989) 'Testing and training of the pelvic floor muscles after childbirth', *Acta Obstetricia et Gynecologica Scandinavica*, 68(4), pp. 301-304.

Joshi, M., Royuela, A., and Zamora, J. (2013) 'Proper analysis in clinical trials: how to report and adjust for missing outcome data', *BJOG: An International Journal of Obstetrics & Gynaecology*, 120(8), pp. 915-919.

Jovanovic, B.D. and Levy, P.S. (1997) 'A look at the rule of three', *The American Statistician*, 51(2), pp. 137-139.

Jovanovic, B.D. and Zalenski, R.J. (1997) 'Safety evaluation and confidence intervals when the number of observed events is small or zero', *Annals of Emergency Medicine*, 30(3), pp. 301-306.

Julious, S.A. (2005) 'Sample size of 12 per group rule of thumb for a pilot study', *Pharmaceutical Statistics*, 4(4), pp. 287-291.

Jung, S.-A., Pretorius, D.H., Padda, B.S., Weinstein, M.M., Nager, C.W., den Boer, D.J., and Mittal, R.K. (2007) 'Vaginal high-pressure zone assessed by dynamic 3-dimensional ultrasound images of the pelvic floor', *American Journal of Obstetrics & Gynecology*, 197(1), pp. 52.e1-52.e7. doi: 10.1016/j.ajog.2007.04.026.

Junginger, B., Baessler, K., Sapsford, R., and Hodges, P.W. (2010) 'Effect of abdominal and pelvic floor tasks on muscle activity, abdominal pressure and bladder neck', *International Urogynecology Journal*, 21(1), pp. 69-77.

Kang, H. (2013) 'The prevention and handling of the missing data', *Korean Journal of Anesthesiology*, 64(5), pp. 402-406.

Kanga Training (no date). Available at: http://www.kangatraining.at/ (Accessed: 26 June 2018).

Kegel, A.H. (1948) 'Progressive resistance exercise in the functional restoration of the perineal muscles', *American Journal of Obstetrics & Gynecology*, 56(2), pp. 238-248.

Kegel, A.H. (1952) 'Sexual functions of the pubococcygeus muscle', *Western Journal of Surgery, Obstetrics, and Gynecology*, 60(10), pp. 521-524.

Kent, H.L. (1991) 'Epidemiology of vaginitis', *American Journal of Obstetrics and Gynecology*, 165(4), pp. 1168-1176.

Kerschan-Schindl, K., Uher, E., Wiesinger, G., Kaider, A., Ebenbichler, G., Nicolakis, P., Kollmitzer, J., Preisinger, E., and Fialka-Moser, V. (2002) 'Reliability of pelvic floor muscle strength measurement in elderly incontinent women', *Neurourology and Urodynamics*, 21(1), pp. 42-47.

Kiene, H. (2001) *Komplementäre Methodenlehre der klinischen Forschung: Cognition-based medicine.* Berlin: Springer. Available at: http://www.ifaemm.de/Abstract/PDFs/CBM_Buch.pdf (Accessed: 11 August 2018).

Kieser Training AG (no date) *Kieser training*. Available at: http://www.kieser-training.com/ (Accessed: 26 June 2018).

Kinouchi, K. and Ohashi, K. (2018) 'Smartphone-based reminder system to promote pelvic floor muscle training for the management of postnatal urinary incontinence: historical control study with propensity score-matched analysis', *PeerJ*, 6, p. e4372. doi: 10.7717/peerj.4372.

Kirk, R.E. (1996) 'Practical significance: a concept whose time has come', *Educational and Psychological Measurement*, 56(5), pp. 746-759.

Knuttgen, H.G. and Kraemer, W.J. (1987) 'Terminology and measurement', *Journal of Applied Sport Science Research*, 1(1), pp. 1-10.

Kondo, A., Yamada, Y., and Niijima, R. (1995) 'Treatment of stress incontinence by vaginal cones: short-and long-term results and predictive parameters', *British Journal of Urology*, 76(4), pp. 464-466.

Köpf-Maier, P. and Wolf-Heidegger, G. (2000) *Wolf-Heidegger's atlas of human anatomy: systemic anatomy, body wall, upper and lower limbs*. 5th edn. Freiburg i.B.: Karger.

Kou, J.-L., Dang, L.-J., and Feng, X.-Q. (2013) 'Clinical study on the treatment of postpartum rehabilitation to improve the pelvic floor function', *Medical Innovation of China 2013*, 10(25), pp. 55-57.

Kraemer, H.C., Mintz, J., Noda, A., Tinklenberg, J., and Yesavage, J.A. (2006) 'Caution regarding the use of pilot studies to guide power calculations for study proposals', *Archives of General Psychiatry*, 63(5), pp. 484-489.

Kraemer, W.J. and Ratamess, N.A. (2004) 'Fundamentals of resistance training: progression and exercise prescription', *Medicine and Science in Sports and Exercise*, 36(4), pp. 674-688.

Kristiansson, P., Samuelsson, E., von Schoultz, B., and Svardsudd, K. (2001) 'Reproductive hormones and stress urinary incontinence in pregnancy', *Acta Obstetricia et Gynecologica Scandinavica*, 80(12), pp. 1125-1130.

Kvale, S. (2013) *Doing interviews*. Los Angeles, CA: Sage Publications Ltd.

LABORIE (2012) *Peritron owner's manual*. Canada: LABORIE.

LABORIE (2018) *Peritron™*. Available at: https://www.laborie.com/products/peritron/ (Accessed: 23 July 2018).

Lagro-Janssen, A., Smits, A., and van Weel, C. (1994) 'Gunstig effect van oefentherapie bij urine-incontinetntie in de huisartspraktijk vooral afhankelijk van therapietrouw en motivatie', *Nederlands Tijdschrift voor Geneeskunde*, 138(25), pp. 1273-1276.

Lagro-Janssen, A.L.M., Debruyne, F.M.J., Smits, A.J.A., and Van Weel, C. (1992) 'The effects of treatment of urinary incontinence in general practice', *Family Practice*, 9(3), pp. 284–289.

Lancaster, G.A. (2015) 'Pilot and feasibility studies come of age!', *Pilot and Feasibility Studies*, 1(1), p. 1. doi: doi:10.1186/2055-5784-1-1.

Lancaster, G.A., Dodd, S., and Williamson, P.R. (2004) 'Design and analysis of pilot studies: recommendations for good practice', *Journal of Evaluation in Clinical Practice*, 10(2), pp. 307-312.

Landon, C., Crofts, C., Smith, A., and Trowbridge, E. (1990) 'Mechanical properties of fascia during pregnancy: a possible factor in the development of stress incontinence of urine', *Contemporary Reviews in Obstetrics and Gynaecology*, 2(4), pp. 40-46.

Lang, T.A. and Altman, D.G. (2014) 'Statistical analyses and methods in the published literature: the SAMPL guidelines*', *Guidelines for Reporting Health Research: A User's Manual.* Chichester: John Wiley & Sons, pp. 264-274.

Lauper, M., Kuhn, A., Gerber, R., Luginbühl, H., and Radlinger, L. (2009) 'Pelvic floor stimulation: what are the good vibrations?', *Neurourology and Urodynamics*, 28(5), pp. 405-410.

Lavin, J., Smith, A., Anderson, J., Grant, M., Buckley, H., Critchley, H., and Hosker, G. (1997) 'The effect of the first pregnancy on the connective tissue of the rectus sheath' [Abstract 27], *Neurourology and Urodynamics*, 16(5), p. 381.

Laycock, J. (2008) 'Concepts of Neuromuscular Rehabilitation and Pelvic Floor Muscle Training', in Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A. & Stanton, S.L. (eds.) *Pelvic floor re-education.* 2nd edn. London: Springer, pp. 177-183.

Laycock, J. and Jerwood, D. (2001) 'Pelvic floor muscle assessment: the PERFECT scheme', *Physiotherapy*, 87(12), pp. 631-642.

Lee, S.-J. and Park, J.-W. (2000) 'Follow-up evaluation of the effect of vaginal delivery on the pelvic floor', *Diseases of the Colon & Rectum*, 43(11), pp. 1550-1555.

Leeman, L.M. and Rogers, R.G. (2012) 'Sex after childbirth: postpartum sexual function', *Obstetrics & Gynecology*, 119(3), pp. 647-655.

Lehmann-Horn, F. (2017) 'Motorische Systeme', in Schmidt, R.F., Lang, F. & Heckmann, M. (eds.) *Physiologie des Menschen.* 31st edn. Berlin: Springer, pp. 127-162.

Leister, N. (2016) E-mail to Claudia Oblasser. 22 July.

LELO (2013a) E-mail to Claudia Oblasser. 8 May.

LELO (2013b) E-mail to Claudia Oblasser. 10 May.

LELO (2013c) *Luna Beads™ user manual & warranty information*. Available at: https://www.lelo.com/luna-beads (Accessed: 25 July 2018).

LELO (2018) *Luna Beads*. Available at: https://www.lelo.com/luna-beads (Accessed: 25 July 2018).

LELO (no date) *Die Kegelübungen verbessern mit LUNA BEADS - Aufwerten der sinnlichen Gesundheit einer Frau.* [Leaflet obtained in Liebenswert Vienna], September 2013.

Lenth, R.V. (2001) 'Some practical guidelines for effective sample size determination', *The American Statistician*, 55(3), pp. 187-193.

Leon, A.C. (2008) 'Implications of clinical trial design on sample size requirements', *Schizophrenia Bulletin*, 34(4), pp. 664-669.

Leon, A.C., Davis, L.L., and Kraemer, H.C. (2011) 'The role and interpretation of pilot studies in clinical research', *Journal of Psychiatric Research*, 45(5), pp. 626-629.

Lesaffre, E. (2008) 'Superiority, equivalence, and non-inferiority trials', *Bulletin of the NYU Hospital for Joint Diseases*, 66(2), pp. 150-154.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P.A., Clarke, M., Devereaux, P.J., Kleijnen, J., and Moher, D. (2009) 'The PRISMA Statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration', *PLoS Medicine* 6(7), p. e1000100. doi: 10.1371/journal.pmed.1000100.

Lien, K.C., Mooney, B., DeLancey, J.O., and Ashton-Miller, J.A. (2004) 'Levator ani muscle stretch induced by simulated vaginal birth', *Obstetrics & Gynecology*, 103(1), pp. 31-40.

Lien, K.C., Morgan, D.M., Delancey, J.O., and Ashton-Miller, J.A. (2005) 'Pudendal nerve stretch during vaginal birth: a 3D computer simulation', *American Journal of Obstetrics & Gynecology*, 192(5), pp. 1669-1676.

Lissworx (2016) *Online Prozentrechner*. Munich: Lissworx. Available at: http://www.prozentrechner.net/ (Accessed: 26 June 2018).

LoBiondo-Wood, G. (2013) 'Introduction to quantitative research', in LoBiondo-Wood, G. & Haber, J. (eds.) *Nursing research: methods and critical appraisal for evidence-based practice*. 8th edn. St. Louis, MO: Elsevier, pp. 163-181.

LoBiondo-Wood, G. and Haber, J. (2013a) *Nursing research: methods and critical appraisal for evidence-based practice*. 8th edn. St. Louis, MO: Mosby.

LoBiondo-Wood, G. and Haber, J. (2013b) 'Reliability and validity', in LoBiondo-Wood, G. & Haber, J. (eds.) *Nursing research: methods and critical appraisal for evidence-based practice*. 8th edn. St. Louis, MO: Elsevier, pp. 289-309.

Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K.E., and Zwarenstein, M. (2015) 'The PRECIS-2 tool: designing trials that are fit for purpose', *BMJ*, 350, p. h2147. doi: 10.1136/bmj.h2147.

Lowenstein, L., Gruenwald, I., Gartman, I., and Vardi, Y. (2010) 'Can stronger pelvic muscle floor improve sexual function?', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 21(5), pp. 553-556.

Luginbuehl, H., Baeyens, J.P., Taeymans, J., Maeder, I.M., Kuhn, A., and Radlinger, L. (2015a) 'Pelvic floor muscle activation and strength components influencing female urinary continence and stress incontinence: a systematic review', *Neurourology and Urodynamcis*, 34(6), pp. 498-506.

Luginbuehl, H., Lehmann, C., Baeyens, J.P., Kuhn, A., and Radlinger, L. (2015b) 'Involuntary reflexive pelvic floor muscle training in addition to standard training versus standard training alone for women with stress urinary incontinence: study protocol for a randomized controlled trial', *Trials*, 16, p. 524. doi: 10.1186/s13063-015-1051-0.

Luginbuehl, H., Lehmann, C., Gerber, R., Kuhn, A., Hilfiker, R., Baeyens, J.P., and Radlinger, L. (2012) 'Continuous versus intermittent stochastic resonance whole body vibration and its effect on pelvic floor muscle activity', *Neurourology and Urodynamics*, 31(5), pp. 683-687.

Lukacz, E.S., Lawrence, J.M., Contreras, R., Nager, C.W., and Luber, K.M. (2006) 'Parity, mode of delivery, and pelvic floor disorders', *Obstetrics & Gynecology*, 107(6), pp. 1253-1260.

Lund Research Ltd (2012) *Non-probability sampling*. Available at: http://dissertation.laerd.com/non-probability-sampling.php#step4 (Accessed: 26 June 2018).

MacArthur, C., Bick, D.E., and Keighley, M.R. (1997) 'Faecal incontinence after childbirth', *British Journal of Obstetrics and Gynaecology*, 104(1), pp. 46-50.

MacArthur, C., Winter, H., Bick, D., Lilford, R., Lancashire, R., Knowles, H., Braunholtz, D., Henderson, C., Belfield, C., and Gee, H. (2003) 'Redesigning postnatal care: a randomised controlled trial of protocol-based midwifery-led care focused on individual women's physical and psychological health needs', *Health Technology Assessment*, 7(37). Available at: https://www.journalslibrary.nihr.ac.uk/hta/hta7370#/abstract (Accessed: 26 June 2018).

Macdonald, S., Magill-Cuerden, J., and Mayes, M. (2011) *Mayes' midwifery: a textbook for midwives*. 14th edn. Edinburgh: Bailliere Tindall.

MacInnes, C.L. (2008) 'Why women leave therapy for stress incontinence', *Nursing Times*, 104(41), pp. 50-53.

Mantle, J. and Versi, E. (1991) 'Physiotherapy for stress urinary incontinence: a national survey', *BMJ*, 302(6779), pp. 753-755.

MAPA GmbH (2017) E-mail to Claudia Oblasser. 31 May.

MAPA GmbH (no date-a) *Neue Kraft nach der Geburt mit dem NUK-Beckenbodentrainer.* [Leaflet obtained at Austrian midwifery assembly], 25 April 2018.

MAPA GmbH (no date-b) *NUK Pelvic Floor Trainer*. Available at: https://www.nuk.de/nuk-pelvic-floor-trainer.html (Accessed: 1 August 2018).

MAPA GmbH (no date-c) *NUK Pelvic Floor Trainer operating instructions*. Available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=10&ved=2ahUKEwjT4rSzj8vcAhWJY1AKHUXVDYIQFjAJegQIARAC&url=https%3A%2F%2Fimages-eu.ssl-images-amazon.com%2Fimages%2FI%2F71f6ZVhR-7S.pdf&usg=AOvVaw2thH8uLduumwiVBYPM2Mb5 (Accessed: 1 August 2018).

Margalith, I., Gillon, G., and Gordon, D. (2004) 'Urinary incontinence in women under 65: quality of life, stress related to incontinence and patterns of seeking health care', *Quality of Life Research*, 13(8), pp. 1381-1390.

Marieb, E.N. (2015) *Essentials of human anatomy & physiology*. 11th edn. Boston, MA: Pearson Education Limited.

Marín, P.J. and Rhea, M.R. (2010) 'Effects of vibration training on muscle power: a meta-analysis', *The Journal of Strength & Conditioning Research*, 24(3), pp. 871-878.

Mark, D.B., Lee, K.L., and Harrell, F.E. (2016) 'Understanding the role of *P* values and hypothesis tests in clinical research', *JAMA Cardiology*, 1(9), pp. 1048-1054.

Marshall, J.E., Raynor, M.D., and Myles, M.F. (2014) *Myles textbook for midwives*. 16th edn. Edinburgh: Saunders, Elsevier.

Marshall, K., Thompson, K.A., Walsh, D.M., and Baxter, G.D. (1998) 'Incidence of urinary incontinence and constipation during pregnancy and postpartum: survey of current findings at the Rotunda Lying-in Hospital', *BJOG: An International Journal of Obstetrics & Gynaecology*, 105(4), pp. 400-402.

Marshall, K., Walsh, D., and Baxter, G. (2002) 'The effect of a first vaginal delivery on the integrity of the pelvic floor musculature', *Clinical Rehabilitation*, 16(7), pp. 795-799.

Martikainen, P., Laaksonen, M., Piha, K., and Lallukka, T. (2007) 'Does survey non-response bias the association between occupational social class and health?', *Scandinavian Journal of Social Medicine*, 35(2), pp. 212-215.

Martinez, C.S., Ferreira, F.V., Castro, A.A., and Gomide, L.B. (2014) 'Women with greater pelvic floor muscle strength have better sexual function', *Acta Obstetricia et Gynecologica Scandinavica*, 93(5), pp. 497-502.

Martins, J., Pato, M., Pires, E., Jorge, R.N., Parente, M., and Mascarenhas, T. (2007) 'Finite element studies of the deformation of the pelvic floor', *Annals of the New York Academy of Sciences*, 1101(1), pp. 316-334.

Mason, L., Glenn, S., Walton, I., and Appleton, C. (1999) 'The prevalence of stress incontinence during pregnancy and following delivery', *Midwifery*, 15(2), pp. 120-128.

Mason, L., Glenn, S., Walton, I., and Hughes, C. (2001a) 'Do women practise pelvic floor exercises during pregnancy or following delivery?', *Physiotherapy*, 87(12), pp. 662-670.

Mason, L., Glenn, S., Walton, I., and Hughes, C. (2001b) 'The instruction in pelvic floor exercises provided to women during pregnancy or following delivery', *Midwifery*, 17(1), pp. 55-64.

Matthews, C.A., Whitehead, W.E., Townsend, M.K., and Grodstein, F. (2013) 'Risk factors for urinary, fecal, or dual incontinence in the Nurses' Health Study', *Obstetrics & Gynecology*, 122(3), pp. 539-545.

Matthews, J.N.S. (2000) *An introduction to randomized controlled clinical trials*. London: Arnold.

Mayo Foundation for Medical Education and Research (2018) *Vaginal atrophy*. Available at: http://www.mayoclinic.org/diseases-conditions/vaginal-atrophy/home/ovc-20200167 (Accessed: 26 June 2018).

McDonald, E., Woolhouse, H., and Brown, S.J. (2015) 'Consultation about sexual health issues in the year after childbirth: a cohort study', *Birth*, 42(4), pp. 354-361.

McDowell, I. (2006) *Measuring health: a guide to rating scales and questionnaires.* Oxford: Oxford University Press. Available at: http://0-www.oxfordscholarship.com.wam.city.ac.uk/view/10.1093/acprof:oso/9780195165678.001.0001/acprof-9780195165678-chapter-10 (Accessed: 28 June 2018).

McKey, P.L. and Dougherty, M.C. (1986) 'The circumvaginal musculature: correlation between pressure and physical assessment', *Nursing Research*, 35(5), pp. 307-309.

McNeilly, A.S. (1979) 'Effects of lactation on fertility', *British Medical Bulletin*, 35(2), pp. 151-154.

'Mechanical vibration' (2003), in *McGraw-Hill dictionary of scientific & technical terms*. 6th edn. Available at: https://encyclopedia2.thefreedictionary.com/mechanical+vibration (Accessed: 10 September 2018).

MedCalc Software bvba (2018) *Intraclass correlation coefficient*. Available at: https://www.medcalc.org/manual/intraclasscorrelation.php (Accessed: 26 June 2018).

medesign I.C. GmbH (2018a) E-mail to Claudia Oblasser. 16 May.

medesign I.C. GmbH (2018b) E-mail to Claudia Oblasser. 9 July.

medesign I.C. GmbH (2018c) *Wie wird Femcon benutzt?* Available at: http://www.myfemcon.de/femcon-benutzung.php (Accessed: 30 July 2018).

medesign I.C. GmbH (no date-a) *Fembowls Vaginalbowls*. Available at: http://en.medesign.de/frontend/shop/showArticle/fembowls-vaginalkugeln/66 (Accessed: 30 July 2018).

medesign I.C. GmbH (no date-b) *Fembowls Vaginalkugeln Gebrauchsanweisung*. Available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjs1-PdsMfcAhWBh6YKHUcTCoAQFjAAegQIABAC&url=https%3A%2F%2Fwww.medesign.de%2Fassets%2Fdownload%2Fproduct%2FBedien_Fembowls_Vaginalkugeln.pdf&usg=AOvVaw2gywKRaey6YNdy24bTcZQn (Accessed: 30 April 2018).

medesign I.C. GmbH (no date-c) *Produktinformationen "Fembowls Vaginalkugeln"*. Available at: https://medesign.de/gynaekologie-pessare/beckenbodentraining/fembowls-vaginalkugeln (Accessed: 30 July 2018).

Medforth, J., Ball, L., Walker, A., Battersby, S., and Stables, S. (2017) *Oxford handbook of midwifery*. 3rd edn. New York, NY: Oxford University Press.

Medgo (1999) *VagaCare vaginal cones*. Available at: https://medgo.com/index.php?main_page=product_info&cPath=36&products_id=35&zenid=2moc3uu7j039qji33k2f0sv096 (Accessed: 26 June 2018).

Medizinische Kontinenzgesellschaft Österreich (2016) *Rat für Blase & Darm Teil 1: Schwangerschaft & Geburt*. Available at:

http://www.kontinenzgesellschaft.at/img/WCW_2016/factsheet_FINAL%20web.pdf (Accessed: 29 June 2018).

Meinert, C.L. (2012) *Clinical trials: design, conduct, and analysis*. 2nd edn. Oxford: Oxford University Press.

Messelink, B., Benson, T., Berghmans, B., Bo, K., Corcos, J., Fowler, C., Laycock, J., Lim, P., van Lunsen, R., and Nijeholt, G.a. (2005) 'Standardization of terminology of pelvic floor muscle function and dysfunction: report from the pelvic floor clinical assessment group of the International Continence Society', *Neurourology and Urodynamics*, 24(4), pp. 374-380.

Meyer, S., Hohlfeld, P., Achtari, C., and De Grandi, P. (2001) 'Pelvic floor education after vaginal delivery', *Obstetrics & Gynecology*, 97(5 Pt 1), pp. 673-677.

Microsoft (2018) *Office 365*. Available at: https://products.office.com/de-at/business/get-latest-office-365-for-your-business-with-2016?omkt=de-AT&WT.mc_id=PS_Google_O365SMB_office%20365&WT.srch=1&omkt=de-AT&WT.mc_id=PS_Google_O365SMB_office%20365&WT.srch=1 (Accessed: 26 June 2018).

Milsom, I., Altman, D., Cartwright, R., Lapitan, M.C., Nelson, R., Sjöström, S., and Tikkinen, K.A.O. (2017) 'Epidemiology of urinary incontinence (UI) and other lower urinary tract symptoms (LUTS), pelvic organ prolapse (POP) and anal (AI) incontinence', in Abrams, P., Cardozo, L., Wagg, A. & Wein, A. (eds.) *Incontinence.* 6th edn. No place: ICUD ICS, pp. 1-141.

Miodrag, A., Castleden, C.M., and Vallance, T.R. (1988) 'Sex hormones and the female urinary tract', *Drugs*, 36(4), pp. 491-504.

Montori, V.M. and Guyatt, G.H. (2001) 'Intention-to-treat principle', *CMAJ: Canadian Medical Association Journal*, 165(10), pp. 1339-1341.

Moore, C.G., Carter, R.E., Nietert, P.J., and Stewart, P.W. (2011) 'Recommendations for planning pilot studies in clinical and translational research', *Clinical and Translational Science*, 4(5), pp. 332-337.

Moore, D.S. and Notz, W.I. (2006) *Statistics: concepts and controversies*. 6th edn. New York, NY: W.H. Freeman and Company.

Moore, K., Dumoulin, C., Bradley, C., Burgio, K., Chambers, T., Hagen, S., Hunter, K., Imamura, M., Thakar, R., Williams, K., and Vale, L. (2013) 'Adult conservative management', in Abrams, P., Cardozo, L., Khoury, S. & Wein, A. (eds.) *Incontinence.* 5th edn. No place: ICUD-EAU, pp. 1101-1227.

Mørkved, S. (2007) 'Evidence for pelvic floor physical therapy for urinary incontinence during pregnancy and after childbirth', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for the pelvic floor: bridging science and clinical practice.* Edinburgh: Churchill Livingstone, pp. 317-336.

Mørkved, S. and Bø, K. (1997) 'The effect of postpartum pelvic floor muscle exercise in the prevention and treatment of urinary incontinence', *International Urogynecology Journal*, 8(4), pp. 217-222.

Mørkved, S. and Bø, K. (1999) 'Prevalence of urinary incontinence during pregnancy and postpartum', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 10(6), pp. 394-398.

Mørkved, S. and Bø, K. (2014) 'Effect of pelvic floor muscle training during pregnancy and after childbirth on prevention and treatment of urinary incontinence: a systematic review', *British Journal of Sports Medicine*, 48(4), pp. 299-310.

Mørkved, S. and Bø, K. (2015) 'Urinary incontinence related to the peripartum period', in Bø, K., Berghmans, B., Mørkved, S. & Van Kampen, M. (eds.) *Evidence-based physical therapy for*

the pelvic floor: bridging science and clinical practice. 2nd edn. Edinburgh: Churchill Livingstone, pp. 208-226.

Mørkved, S., Salvesen, K.A., Bø, K., and Eik-Nes, S. (2004) 'Pelvic floor muscle strength and thickness in continent and incontinent nulliparous pregnant women', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 15(6), pp. 384-390.

'Muscular hypertrophy' (2007), in *Saunders comprehensive veterinary dictionary*. 3rd edn. Available at: https://medical-dictionary.thefreedictionary.com/muscular+hypertrophy (Accessed: 29 June 2018).

Myer, E.N.B., Roem, J.L., Lovejoy, D.A., Abernethy, M.G., Blomquist, J.L., and Handa, V.L. (2018) 'Longitudinal changes in pelvic floor muscle strength among parous women', *American Journal of Obstetrics and Gynecology*. doi: 10.1016/j.ajog.2018.06.003.

Mylonas, I. and Friese, K. (2010) *Infektionen in Gynäkologie und Geburtshilfe*. München: Urban & Fischer, Elsevier.

National Collaborating Centre for Women's and Children's Health (2013) *Urinary incontinence in women: the management of urinary incontinence in women*. Available at: https://www.nice.org.uk/guidance/cg171/evidence/full-guideline-updated-november-2015-pdf-191581165 (Accessed: 29 June 2018).

National Institute for Health Research (2017) *NIHR Research for Patient Benefit (RfPB) Programme: guidance on applying for feasibility studies*. Available at: https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/research-for-patient-benefit/ (Accessed: 19 July 2018).

National Institute for Health Research (no date-a) *Feasibility and pilot studies*. Available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwi7mYulkpncAhUObFAKHbDMAp8QFggtMAA&url=https%3A%2F%2Fwww.nihr.ac.uk%2Ffunding-and-support%2Fdocuments%2Ffunding-for-research-studies%2Fresearch-programmes%2FPGfAR%2FCCF-PGfAR-Feasibility-and-Pilot-studies.pdf&usg=AOvVaw3W6W4kUv1eh4Jcl44yUkty (Accessed: 30 September 2018).

National Institute for Health Research (no date-b) *Patient and public involvement in health and social care research: a handbook for researchers*. Available at: https://www.nihr.ac.uk/about-us/CCF/funding/how-we-can-help-you/RDS-PPI-Handbook-2014-v8-FINAL.pdf (Accessed: 23 July 2018).

National Institute for Health Research CLAHRC South West Peninsula (2018) *APPEAL (Antenatal Preventative Pelvic floor Exercises And Localisation) - delivery of pelvic floor muscle exercise education for women during pregnancy*. Available at: http://clahrc-peninsula.nihr.ac.uk/research/appeal-antenatal-preventative-pelvic-floor-exercises-and-localisation-delivery-of-pelvic-floor-muscle-exercise-education-for-wom (Accessed: 19 July 2018).

National Institute for Health Research Clinical Research Network (NIHR CRN) (2016) *Good Clinical Practice (GCP) reference guide*. Available at: https://www.nihr.ac.uk/our-faculty/clinical-research-staff/learning-and-development/national-directory/good-clinical-practice/ (Accessed: 24 July 2018).

National Institutes of Health (2016) *Definitions of criteria and considerations for research project grant (RPG/X01/R01/R03/R21/R33/R34) critiques*. Available at: https://grants.nih.gov/grants/peer/critiques/rpg_D.htm (Accessed: 23 July 2018).

National Institutes of Health (no date) *NIH guide to grants and contracts*. Available at: https://grants.nih.gov/grants/funding/funding_program.htm (Accessed: 6 August 2018).

National Statistics (2003) *Ethnic group statistics: a guide for the collection and classification of ethnicity data*. No place: HMSO. Available at:

https://www.google.at/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjs4IKh4vHJAhVJ0RQKHSt0BlQQFggcMAA&url=http%3A%2F%2Fwww.ons.gov.uk%2Fons%2Fguide-method%2Fmeasuring-equality%2Fequality%2Fmeasuring-equality--a-guide%2Fethnic-group-statistics--a-guide-for-the-collection-and-classification-of-ethnicity-data.pdf&usg=AFQjCNGpS_4zu1AeVJAsp4OlqapiXTWHbQ&cad=rjt (Accessed: 29 June 2018).

Neels, H., Mortiers, X., de Graaf, S., Tjalma, W.A.A., De Wachter, S., and Vermandel, A. (2017) 'Vaginal wind: A literature review', *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 214, pp. 97-103.

NETSCC (2016) *Feasibility and pilot studies: which programme should I apply to?* Available at: https://www.nihr.ac.uk/glossary?letter=F&postcategory=-1 (Accessed: 23 July 2018).

Newington, L. and Metcalfe, A. (2014) 'Factors influencing recruitment to research: qualitative study of the experiences and perceptions of research teams', *BMC Medical Research Methodology*, 14(1), p. 10. doi: 10.1186/1471-2288-14-10.

NHS Choices Team (2016) *Clinical trials*. Available at: http://www.nhs.uk/Conditions/Clinical-trials/Pages/Introduction.aspx (Accessed: 23 July 2018).

Nielsen, C.A., Sigsgaard, I., Olsen, M., Tolstrup, M., Danneskiold-Samsoee, B., and Bock, J.E. (1988) 'Trainability of the pelvic floor. A prospective study during pregnancy and after delivery', *Acta Obstetricia et Gynecologica Scandinavica*, 67(5), pp. 437-440.

Nolan, M. (2001) *Professionelle Geburtsvorbereitung: Geburtsvorbereitungskurse erfolgreich planen, durchführen und bewerten*. Bern: Hans Huber.

Noordzij, M., Tripepi, G., Dekker, F.W., Zoccali, C., Tanck, M.W., and Jager, K.J. (2010) 'Sample size calculations: basic principles and common pitfalls', *Nephrology Dialysis Transplantation*, 25(5). doi: 10.1093/ndt/gfp732.

Norton, P. and Baker, J. (1990) 'Randomized prospective trial of vaginal cones vs. Kegel exercises in postpartum primiparous women' [Abstract], *Neurourology and Urodynamics*, 9(4), pp. 434-435.

O'Boyle, A., O'Boyle, J., Ricks, R., Patience, T., Calhoun, B., and Davis, G. (2003) 'The natural history of pelvic organ support in pregnancy', *International Urogynecology Journal*, 14(1), pp. 46-49.

O'Boyle, A.L., Woodman, P.J., O'Boyle, J.D., Davis, G.D., and Swift, S.E. (2002) 'Pelvic organ support in nulliparous pregnant and nonpregnant women: a case control study', *American Journal of Obstetrics & Gynecology*, 187(1), pp. 99-102.

O'Cathain, A., Thomas, K.J., Drabble, S.J., Rudolph, A., Goode, J., and Hewison, J. (2014) 'Maximising the value of combining qualitative research and randomised controlled trials in health research: the QUAlitative Research in Trials (QUART) study - a mixed methods study', *Health Technology Assessment*, 18(38). doi: 10.3310/hta18380.

O'Connor, D., Green, S., and Higgins, J.P.T. (2011) 'Chapter 5: Defining the review question and developing criteria for including studies', in Higgins, J.P.T. & Green, S. (eds.) *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. No place: The Cochrane Collaboration. Available at: http://handbook-5-1.cochrane.org/ (Accessed: 11 August 2017).

O'Hagan, A., Stevens, J.W., and Campbell, M.J. (2005) 'Assurance in clinical trial design', *Pharmaceutical Statistics*, 4(3), pp. 187-201.

O'Reilly, R., Peters, K., Beale, B., and Jackson, D. (2009) 'Women's experiences of recovery from childbirth: focus on pelvis problems that extend beyond the puerperium', *Journal of Clinical Nursing*, 18(14), pp. 2013-2019.

Oblasser, C. (2015) *Vibrating vaginal balls after childbirth*. Available at: https://www.clinicaltrials.gov/ct2/show/NCT02355327 (Accessed: 26 June 2018).

Oblasser, C., Christie, J., and McCourt, C. (2014a) 'Vaginal cones or balls to improve pelvic floor muscle performance and urinary continence in women post partum: a quantitative systematic review', *PROSPERO* (CRD42014006165). Available at: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014006165#.VHDpHcl0CCc (Accessed: 17 July 2018).

Oblasser, C., Christie, J., and McCourt, C. (2014b) 'Vaginal cones or balls to improve pelvic floor muscle performance and urinary continence in women postpartum: a quantitative systematic review and meta-analysis protocol', *Journal of Advanced Nursing*, 71(4), pp. 933-941.

Oblasser, C., Christie, J., and McCourt, C. (2015) 'Vaginal cones or balls to improve pelvic floor muscle performance and urinary continence in women post partum: a quantitative systematic review', *Midwifery*, 31(11), pp. 1017-1025.

Oblasser, C., McCourt, C., Hanzal, E., and Christie, J. (2016) 'Vibrating vaginal balls to improve pelvic floor muscle performance in women after childbirth: a protocol for a randomised controlled feasibility trial', *Journal of Advanced Nursing*, 72(4), pp. 900-914.

'Observer variation' (2008), in *Mosby's dental dictionary*. 2nd edn. Available at: http://medical-dictionary.thefreedictionary.com/observer+variation (Accessed: 26 June 2018).

OCEBM Levels of Evidence Working Group (2011) *The Oxford 2011 levels of evidence*. Oxford Centre for Evidence-Based Medicine. Available at: http://www.cebm.net/index.aspx?o=5653 (Accessed: 23 July 2018).

Oppenheim, A.N. (2000) *Questionnaire design, interviewing and attitude measurement*. 2nd edn. London: Pinter Publishers.

Orsmond, G.I. and Cohn, E.S. (2015) 'The distinctive features of a feasibility study: objectives and guiding questions', *OTJR: Occupation, Participation and Health*, 35(3), pp. 169-177.

Österreichisches Hebammengremium (2018a) E-mail to Claudia Oblasser. 19 May 2018.

Österreichisches Hebammengremium (2018b) E-mail to Claudia Oblasser. 22 July 2018.

Österreichisches Hebammengremium (2018c) E-mail to Claudia Oblasser. 19 May.

Österreichisches Hebammengremium (2018d) E-mail to Claudia Oblasser. 22 March.

Otcenasek, M., Krofta, L., Baca, V., Grill, R., Kucera, E., Herman, H., Vasicka, I., Drahonovsky, J., and Feyereisl, J. (2007) 'Bilateral avulsion of the puborectal muscle: magnetic resonance imaging-based three-dimensional reconstruction and comparison with a model of a healthy nulliparous woman', *Ultrasound in Obstetrics & Gynecology*, 29(6), pp. 692-696.

Oxford University Press (2018) *Reflectory*. Available at: https://en.oxforddictionaries.com/definition/reflectory (Accessed: 17 September 2018).

Palmezoni, V.P., Santos, M.D., Pereira, J.M., Bernardes, B.T., Pereira-Baldon, V.S., and Resende, A. (2017) 'Pelvic floor muscle strength in primigravidae and non-pregnant nulliparous women: a comparative study', *International Urogynecology Journal*, 28(1), pp. 131-137.

Parmar, M.K., Ungerleider, R.S., and Simon, R. (1996) 'Assessing whether to perform a confirmatory randomized clinical trial', *Journal of the National Cancer Institute*, 88(22), pp. 1645-1651.

Partridge, A.H. and Winer, E.P. (2002) 'Informing clinical trial participants about study results', *JAMA*, 288(3), pp. 363-365.

Patterson Medical (no date) *Aquaflex®*. Available at: http://www.neenpelvichealth.com/uploads/products/aquaflex/neen-aquaflex116pp.pdf (Accessed: 24 July 2018).

Paulsen, F. and Waschke, J. (eds.) (2013) *Sobotta Atlas der Anatomie des Menschen: Becken und Retroperitonealraum.* 23rd edn. München: Elsevier, Urban & Fischer.

Pearson, S. (2015) *Inner thigh, pelvic floor and core exercise on Power Plates with Pear Projects*. Available at: https://www.youtube.com/watch?v=e-rCgYr6iLA (Accessed: 26 June 2018).

Peattie, A. and Plevnik, S. (1988) 'Cones versus physiotherapy as conservative management of genuine stress incontinence ' [Abstract 72] (from the Proceedings of the 18th annual meeting of the International Continence Society, Oslo, 1-3 September 1988), *Neurourology and Urodynamics*, 7(3), pp. 255-256.

Peattie, A.B., Plevnik, S., and Stanton, S.L. (1988) 'Vaginal cones: a conservative method of treating genuine stress incontinence', *British Journal of Obstetrics and Gynaecology*, 95(10), pp. 1049-1053.

Pelvic, Obstetric and Gynaecological Physiotherapy (2017a) E-mail to Claudia Oblasser. 17 May.

Pelvic, Obstetric and Gynaecological Physiotherapy (2017b) *Fit for the future: essential exercises and advice after childbirth.* Available at: http://pogp.csp.org.uk/publications/fit-future (Accessed: 31 July 2018).

Perera, R., Heneghan, C., and Yudkin, P. (2007) 'Graphical method for depicting randomised trials of complex interventions', *BMJ*, 334(7585), pp. 127-129.

Perucchini, D. and DeLancey, J.O.L. (2008) 'Functional anatomy of the pelvic floor and lower urinary tract', in Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A. & Stanton, S.L. (eds.) *Pelvic floor re-education: principles and practice.* 2nd edn. London: Springer, pp. 3-21.

Peschers, U., Schaer, G., Anthuber, C., Delancey, J.O., and Schuessler, B. (1996) 'Changes in vesical neck mobility following vaginal delivery', *Obstetrics & Gynecology*, 88(6), pp. 1001-1006.

Peschers, U.M., Gingelmaier, A., Jundt, K., Leib, B., and Dimpfl, T. (2001) 'Evaluation of pelvic floor muscle strength using four different techniques', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 12(1), pp. 27-30.

Peschers, U.M., Schaer, G.N., DeLancey, J.O., and Schuessler, B. (1997) 'Levator ani function before and after childbirth', *British Journal of Obstetrics and Gynaecology*, 104(9), pp. 1004-1008.

Petrie, A. and Sabin, C. (2009) *Medical statistics at a glance*. 3rd edn. Chichester: Wiley-Blackwell.

Pezeshk, H. (2003) 'Bayesian techniques for sample size determination in clinical trials: a short review', *Statistical Methods in Medical Research*, 12(6), pp. 489-504.

Pilates Foundation (2018) *Pilates*. Available at: http://www.pilatesfoundation.com/pilates/ (Accessed: 26 June 2018).

'Placebo effect' (2007), in *The American heritage® medical dictionary*. Available at: https://medical-dictionary.thefreedictionary.com/placebo+effect (Accessed: 26 June 2018).

Plevnik, S. (1985) 'New method for testing and strengthening of pelvic floor muscles' [Abstract] (from the Proceedings of the 15th annual meeting of the ICS 1985, London, 3-6 September), pp. 267-268.

Pocock, S.J. (2003) 'The pros and cons of noninferiority trials', *Fundamental & Clinical Pharmacology*, 17(4), pp. 483-490.

Pocock, S.J. and Stone, G.W. (2016) 'The primary outcome is positive—is that good enough?', *New England Journal of Medicine*, 375(10), pp. 971-979.

Porta-Roda, O., Simo Gonzalez, M., Reula Blasco, M.C., Diaz Lopez, M.A., Diaz Bellido, P., Vara Paniagua, J., Sobrado Lozano, P., and Muñoz Garrido, F. (2013) 'Use of a vaginal spheres device in the conservative treatment of stress urinary incontinence: a randomized controlled trial' [Poster], *International Continence Society annual meeting*, Barcelona, 26-30 August. Available at: http://www.ics.org/Abstracts/Publish/180/000099_poster.pdf (Accessed: 29 June 2018).

Porta-Roda, O., Vara-Paniagua, J., Diaz-Lopez, M.A., Sobrado-Lozano, P., Simo-Gonzalez, M., Diaz-Bellido, P., Reula-Blasco, M.C., and Munoz-Garrido, F. (2015) 'Effect of vaginal spheres and pelvic floor muscle training in women with urinary incontinence: a randomized, controlled trial', *Neurourology and Urodynamcis*, 34(6), pp. 533-538.

Powerfilm GbR (no date) *Informationen zu den Smartballs*. Available at: http://www.rueckbildungsgymnastik.net/forum/thread.php?threadid=219 (Accessed: 23 September 2018).

Prashar, S., Simons, A., Bryant, C., Dowell, C., and Moore, K.H. (2000) 'Attitudes to vaginal/urethral touching and device placement in women with urinary incontinence', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 11(1), pp. 4-8.

Propagator and NHS (2018) *Squeezy*. Available at: http://www.squeezyapp.co.uk/ (Accessed: 26 June 2018).

'Proprioception' (2012), in *Segen's medical dictionary*. Available at: http://medical-dictionary.thefreedictionary.com/proprioception (Accessed: 29 June 2018).

Pschyrembel, W. (2013) *Pschyrembel Klinisches Wörterbuch 2014*. 265th edn. Berlin: Walter de Gruyter.

PT Direct (2018) *PelvicToner™*. Available at: http://www.pelvictoner.co.uk/ (Accessed: 26 June 2018).

Quigley, E.M. (2007) 'Impact of pregnancy and parturition on the anal sphincters and pelvic floor', *Best Practice & Research: Clinical Gastroenterology*, 21(5), pp. 879-891.

Rahmani, N. and Mohseni-Bandpei, M.A. (2011) 'Application of perineometer in the assessment of pelvic floor muscle strength and endurance: a reliability study', *Journal of Bodywork & Movement Therapies*, 15(2), pp. 209-214.

Ratamess, N.A. (2012) *ACSM's foundations of strength training and conditioning*. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins.

'Reflex' (2014), in *The American heritage® student science dictionary*. 2nd edn. Available at: https://www.thefreedictionary.com/reflex (Accessed: 18 September 2018).

Reilly, E.T.C., Freeman, R.M., Waterfield, M.R., Waterfield, A.E., Steggles, P., and Pedlar, F. (2002) 'Prevention of postpartum stress incontinence in primigravidae with increased bladder neck mobility: a randomised controlled trial of antenatal pelvic floor exercises', *BJOG: An International Journal of Obstetrics & Gynaecology*, 109(1), pp. 68-76.

René Descartes (2018) *René Descartes > Quotes > Quotable Quote.* Available at: https://www.goodreads.com/quotes/484961-dubium-sapientiae-initium-doubt-is-the-origin-of-wisdom (Accessed: 26 June 2018).

ResearchNet (2018) *Funding opportunity details*. Available at: https://www.researchnet-recherchenet.ca/rnr16/vwOpprtntyDtls.do?prog=2388&view=search&terms=pilot&org=CIHR&type=EXACT&resultCount=25%20%20%20http://www.cihr-irsc.gc.ca/e/45336.html (Accessed: 23 July 2018).

Ribeiro, D.C., Sole, G., Abbott, J.H., and Milosavljevic, S. (2014) 'The effectiveness of a lumbopelvic monitor and feedback device to change postural behavior: a feasibility randomized controlled trial', *Journal of Orthopaedic & Sports Physical Therapy*, 44(9), pp. 702-711.

Rittweger, J., Beller, G., and Felsenberg, D. (2000) 'Acute physiological effects of exhaustive whole-body vibration exercise in man', *Clinical Physiology*, 20(2), pp. 134-142.

Robinson, D. and Cardozo, L. (2008) 'Hormonal influences on the female genital and lower urinary tract', in Raz, S. & Rodriguez, L.V. (eds.) *Female Urology.* 3rd edn. Philadelphia, PA: Saunders Elsevier, pp. 59-68.

Rochera, M.B., Andreu, C.S., Madrid, Y.C., Bouallalene, K., and Pau, E.M. (2017) 'Physiotherapy as a way to maintain vaginal health during menopause', *Advances in Sexual Medicine*, 7(2), pp. 97-104.

Rodrigues, M.P., Paiva, L.L., Ramos, J.G.L., and Ferla, L. (2018) 'Vibratory perineal stimulation for the treatment of female stress urinary incontinence: a systematic review', *International Urogynecology Journal*, 29(4), pp. 555-562.

Römer, T., Schleußner, E., and Straube, W. (eds.) (2012) *Pschyrembel Gynäkologie und Geburtshilfe.* 3rd edn. Berlin: de Gruyter.

Rortveit, G., Brown, J.S., Thom, D.H., Van Den Eeden, S.K., Creasman, J.M., and Subak, L.L. (2007) 'Symptomatic pelvic organ prolapse: prevalence and risk factors in a population-based, racially diverse cohort', *Obstetrics & Gynecology*, 109(6), pp. 1396-1403.

Rosenbaum, T.Y. (2007) 'Pelvic floor involvement in male and female sexual dysfunction and the role of pelvic floor rehabilitation in treatment: a literature review', *The Journal of Sexual Medicine*, 4(1), pp. 4-13.

Royal College of Midwives (2009) *How to… teach pelvic floor muscle exercises*. Available at: https://www.rcm.org.uk/news-views-and-analysis/analysis/how-to%E2%80%A6-teach-pelvic-floor-muscle-exercises (Accessed: 21 August 2018).

Royal College of Midwives (2018) *A third of mothers don't do pelvic floor exercises*. Available at: https://www.rcm.org.uk/news-views-and-analysis/news/a-third-of-mothers-don%E2%80%99t-do-pelvic-floor-exercises (Accessed: 21 August 2018).

Royal College of Midwives (no date) *Peri-partum pelvic floor muscle exercises*. Available at: http://www.ilearn.rcm.org.uk/enrol/index.php?id=58 (Accessed: 14 July 2018).

Sackett, D.L. (1979) 'Bias in analytic research', *Journal of Chronic Diseases*, 32(1-2), pp. 51-63.

Salinas Casado, J., Virseda Chamorro, M., Esteban Fuertes, M., and Grupo RECOVA (1999) 'Eficacia terapeutica de los conos vaginales en la rehabilitacion de la musculatura perineal', *Archivos Españoles de Urología*, 52(1), pp. 53-60.

Salvatore, S., Rademakers, K., DeLancey, J., Igawa, Y., Koelbl, H., Laterza, R.M., Serati, M., Sultan, A., Sievert, K.D., and Lowry, A. (2017) 'Pathophysiology of urinary incontinence, faecal incontinence and pelvic organ prolapse', in Abrams, P., Cardozo, L., Wagg, A. & Wein, A. (eds.) *Incontinence.* 6th edn. No place: ICUD ICS, pp. 361-495.

Sampselle, C.M. (1990) 'Changes in pelvic muscle strength and stress urinary incontinence associated with childbirth', *JOGNN: Journal of Obstetric, Gynecologic & Neonatal Nursing*, 19(5), pp. 371-377.

Sangsawang, B. and Sangsawang, N. (2013) 'Stress urinary incontinence in pregnant women: a review of prevalence, pathophysiology, and treatment', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 24(6), pp. 901-912.

Sankey, S.S. and Weissfeld, L.A. (1998) 'A study of the effect of dichotomizing ordinal data upon modeling', *Communications in Statistics - Simulation and Computation*, 27(4), pp. 871-887.

Sapsford, R., Hodges, P., Richardson, C., Cooper, D., Markwell, S., and Jull, G. (2001) 'Co-activation of the abdominal and pelvic floor muscles during voluntary exercises', *Neurourology and Urodynamics*, 20(1), pp. 31-42.

Sapsford, R.R. and Hodges, P.W. (2001) 'Contraction of the pelvic floor muscles during abdominal maneuvers', *Archives of Physical Medicine and Rehabilitation*, 82(8), pp. 1081-1088.

Schildbach, S. (2005) 'Beckenbodentraining mit Geisha-Kugeln', *Psyche und Soma*, 27(11), pp. 28-30.

Schmidt, C.O. and Kohlmann, T. (2008) 'When to use the odds ratio or the relative risk?', *International Journal of Public Health*, 53(3), pp. 165-167.

Schmidt, R.F. and Schaible, H.-G. (2006) *Neuro- und Sinnesphysiologie*. 5th edn. Heidelberg: Springer.

Schoenfeld, D. (1980) 'Statistical considerations for pilot studies', *International Journal of Radiation Oncolocy Biology Physics*, 6(3), pp. 371-374.

Schulz, K.F., Altman, D.G., and Moher, D. (2010) 'CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials', *BMC Medicine*, 8, p. 18. doi: 10.1186/1741-7015-8-18 (Accessed: 3 February 2019).

Schulz, K.F. and Grimes, D.A. (2002) 'Unequal group sizes in randomised trials: guarding against guessing', *The Lancet*, 359(9310), pp. 966-970.

Schulz, K.F. and Grimes, D.A. (2005) 'Sample size calculations in randomised trials: mandatory and mystical', *The Lancet*, 365(9467), pp. 1348-1353.

Schwartz, D. and Lellouch, J. (1967) 'Explanatory and pragmatic attitudes in therapeutical trials', *Journal of Chronic Diseases*, 20(8), pp. 637-648.

Schwertner-Tiepelmann, N., Thakar, R., Sultan, A.H., and Tunn, R. (2012) 'Obstetric levator ani muscle injuries: current status', *Ultrasound in Obstetrics & Gynecology*, 39(4), pp. 372-383.

Sealed Envelope Ltd (2018) *Create a blocked randomisation list*. Available at: https://www.sealedenvelope.com/simple-randomiser/v1/lists (Accessed: 26 June 2018).

Senat, M.V., Sentilhes, L., Battut, A., Benhamou, D., Bydlowski, S., Chantry, A., Deffieux, X., Diers, F., Doret, M., Ducroux-Schouwey, C., Fuchs, F., Gascoin, G., Lebot, C., Marcellin, L., Plu-Bureau, G., Raccah-Tebeka, B., Simon, E., Breart, G., and Marpeau, L. (2016) 'Postpartum practice: guidelines for clinical practice from the French College of Gynaecologists and Obstetricians (CNGOF)', *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 202, pp. 1-8.

Seo, J.T., Yoon, H., and Kim, Y.H. (2004) 'A randomized prospective study comparing new vaginal cone and FES-Biofeedback', *Yonsei Medical Journal*, 45(5), pp. 879-884.

Shafik, A. (1979) 'A new concept of the anatomy of the anal sphincter mechanism and the physiology of defecation. VIII. Levator hiatus and tunnel: anatomy and function', *Diseases of the Colon and Rectum*, 22(8), pp. 539-549.

Shanyinde, M., Pickering, R.M., and Weatherall, M. (2011) 'Questions asked and answered in pilot and feasibility randomized controlled trials', *BMC Medical Research Methodology*, 11, p. 117. doi: 10.1186/1471-2288-11-117.

Shek, K.L., Chantarasorn, V., Langer, S., and Dietz, H.P. (2012a) 'Does levator trauma "heal"?', *Ultrasound in Obstetrics & Gynecology*, 40(5), pp. 570-575.

Shek, K.L., Chantarasorn, V., Langer, S., Phipps, H., and Dietz, H.P. (2011) 'Does the Epi-No Birth Trainer reduce levator trauma? A randomised controlled trial', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 22(12), pp. 1521-1528.

Shek, K.L. and Dietz, H.P. (2009) 'The effect of childbirth on hiatal dimensions', *Obstetrics & Gynecology*, 113(6), pp. 1272-1278.

Shek, K.L., Kruger, J., and Dietz, H.P. (2012b) 'The effect of pregnancy on hiatal dimensions and urethral mobility: an observational study', *International Urogynecology Journal*, 23(11), pp. 1561-1567.

Sherwood, L. (2016) *Human physiology: from cells to systems*. 9th edn. Boston, MA: Cengage Learning.

Shizgal, P.B. and Hyman, S.E. (2013) 'Homeostasis, motivation, and addictive states', in Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S.A. & Hudspeth, A.J. (eds.) *Principles of neural science.* 5th edn. New York, NY: McGraw-Hill Companies, Inc., pp. 1095-1115.

Shobeiri, S.A., Chesson, R.R., and Gasser, R.F. (2008) 'The internal innervation and morphology of the human female levator ani muscle', *American Journal of Obstetrics & Gynecology*, 199(6), pp. 686.e1-686.e6.

Sigurdardottir, T., Steingrimsdottir, T., Arnason, A., and Bø, K. (2011) 'Pelvic floor muscle function before and after first childbirth', *International Urogynecology Journal*, 22(12), pp. 1497-1503.

Silbernagl, S. and Despopoulos, A. (2012) *Taschenatlas Physiologie*. 8th edn. Stuttgart: Thieme.

Sim, J. and Lewis, M. (2012) 'The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency', *Journal of Clinical Epidemiology*, 65(3), pp. 301-308.

Simon, S. (2010) *Pilot study*. Available at: http://www.pmean.com/99/pilot.html (Accessed: 29 June 2018).

Sleep, J. and Grant, A. (1987) 'Pelvic floor exercises in postnatal care', *Midwifery*, 3(4), pp. 158-164.

Slieker-ten Hove, M.C.P., Pool-Goudzwaard, A.L., Eijkemans, M.J., Steegers-Theunissen, R.P., Burger, C.W., and Vierhout, M.E. (2009) 'Pelvic floor muscle function in a general female population in relation with age and parity and the relation between voluntary and involuntary contractions of the pelvic floor musculature', *International Urogynecology Journal*, 20(12), pp. 1497-1504.

Slieker-ten Hove, M., Pool-Goudzwaard, A., Eijkemans, M., Steegers-Theunissen, R., Burger, C., and Vierhout, M. (2009) 'Face validity and reliability of the first digital assessment scheme of pelvic floor muscle function conform the new standardized terminology of the International Continence Society', *Neurourology and Urodynamics*, 28(4), pp. 295-300.

Snooks, S.J., Setchell, M., Swash, M., and Henry, M.M. (1984) 'Injury to innervation of pelvic floor sphincter musculature in childbirth', *The Lancet*, 2(8402), pp. 546-550.

Sobel, J.D., Faro, S., Force, R.W., Foxman, B., Ledger, W.J., Nyirjesy, P.R., Reed, B.D., and Summers, P.R. (1998) 'Vulvovaginal candidiasis: epidemiologic, diagnostic, and therapeutic considerations', *American Journal of Obstetrics & Gynecology*, 178(2), pp. 203-211.

Solans-Domenech, M., Sanchez, E., and Espuna-Pons, M. (2010) 'Urinary and anal incontinence during pregnancy and postpartum: incidence, severity, and risk factors', *Obstetrics & Gynecology*, 115(3), pp. 618-628.

Stafne, S., Salvesen, K., Romundstad, P., Torjusen, I., and Mørkved, S. (2012) 'Does regular exercise including pelvic floor muscle training prevent urinary and anal incontinence during pregnancy?

A randomised controlled trial', *BJOG: An International Journal of Obstetrics & Gynaecology*, 119(10), pp. 1270-1280.

Stallard, N. (2012) 'Optimal sample sizes for phase II clinical trials and pilot studies', *Statistics in Medicine*, 31(11-12), pp. 1031-1042.

Staniszewska, S., Brett, J., Simera, I., Seers, K., Mockford, C., Goodlad, S., Altman, D., Moher, D., Barber, R., Denegri, S., Entwistle, A., Littlejohns, P., Morris, C., Suleman, R., Thomas, V., and Tysall, C. (2017) 'GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research', *Research Involvement and Engagement*, 3, p. 13. doi: 10.1186/s40900-017-0062-2.

Stanley, K. (2007) 'Design of randomized controlled trials', *Circulation*, 115(9), pp. 1164-1169.

Stanton, S., Kerr-Wilson, R., and Harris, V.G. (1980) 'The incidence of urological symptoms in normal pregnancy', *BJOG: An International Journal of Obstetrics & Gynaecology*, 87(10), pp. 897-900.

Statistics How To (2018) *Cronbach's Alpha: simple definition, use and interpretation* Available at: http://www.statisticshowto.com/cronbachs-alpha-spss/ (Accessed: 23 July 2018).

Statistik Austria (2015) Telephone call with Claudia Oblasser. No date.

Statistik Austria (2017a) *Lebendgeborene (mit inländischem Geburtsort) 2016 nach Geburtsgewicht und weiteren Merkmalen*. Available at: https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/geborene/medizinische_und_sozialmedizinische_merkmale/index.html (Accessed: 25 January 2018).

Statistik Austria (2017b) *Totgeborene und späte Fetalsterblichkeit seit 2000 nach Wohnbundesland der Mutter*. Available at: https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/geborene/medizinische_und_sozialmedizinische_merkmale/index.html (Accessed: 25 January 2018).

Statistik Austria (2018a) E-mail to Claudia Oblasser. 16 January 2018.

Statistik Austria (2018b) *Jahrbuch der Gesundheitsstatistik 2016*. Wien: Statistik Austria.

StatPages.net (2009) *Exact binomial and poisson confidence intervals*. Available at: http://statpages.info/confint.html#Binomial (Accessed: 29 June 2018).

Stats Make Me Cry Consulting (no date) *Confusing stats terms explained: internal consistency*. Available at: http://www.statsmakemecry.com/smmctheblog/confusing-stats-terms-explained-internal-consistency.html (Accessed: 23 July 2018).

StatTrek.com (2018) *F distribution calculator*. Available at: http://stattrek.com/online-calculator/f-distribution.aspx (Accessed: 26 June 2018).

Steiner, R. (2003) 'Exzentrische Muskelarbeit – die unbekannte Seite unserer Bewegungen', *Therapeutische Umschau*, 60(7), pp. 425-429.

Stengel, D. (ed.) (2010) *Statistik und Aufbereitung klinischer Daten*. Stuttgart: Thieme.

Stewart, E. (2006) 'The roles of the continence nurse specialist', in Cardozo, L. & Staskin, D. (eds.) *Textbook of female urology and urogynecology.* 2nd edn. Oxon: Informa Healthcare, pp. 81-90.

Stothers, L. (2002) 'A randomzed controlled trial to evaluate intrapartum pelvic floor exercise as a method of preventing urinary incontinence' [Abstract] (from the Proceedings of the American Urological Association annual meeting, Orlando, 25-30 May 2002), *Journal of Urology*, 167(4), p. 106.

Sultan, A.H. and Abulaffi, A.M. (2008) 'Anal incontinence, constipation, and obstructed defacation', in Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A. & Stanton, S.L. (eds.) *Pelvic floor re-education: principles and practice.* 2nd edn. London: Springer, pp. 75-82.

Sultan, A.H., Kamm, M.A., Hudson, C.N., Thomas, J.M., and Bartram, C.I. (1993) 'Anal-sphincter disruption during vaginal delivery', *New England Journal of Medicine*, 329(26), pp. 1905-1911.

Svabik, K., Shek, K.L., and Dietz, H.P. (2009) 'How much does the levator hiatus have to stretch during childbirth?', *BJOG: An International Journal of Obstetrics & Gynaecology*, 116(12), pp. 1657-1662.

Sze, E.H., Sherard, G.B., 3rd, and Dolezal, J.M. (2002) 'Pregnancy, labor, delivery, and pelvic organ prolapse', *Obstetrics & Gynecology*, 100(5 Pt 1), pp. 981-986.

Tacke, L. and Stüwe, M. (2013) *Wochenbett-und Rückbildungsgymnastik*. Stuttgart: Georg Thieme.

Teare, M.D., Dimairo, M., Shephard, N., Hayman, A., Whitehead, A., and Walters, S.J. (2014) 'Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study', *Trials*, 15(1), p. 264. doi: 10.1186/1745-6215-15-264.

Tennfjord, M.K., Hilde, G., Staer-Jensen, J., Siafarikas, F., Engh, M.E., and Bo, K. (2015) 'Coital incontinence and vaginal symptoms and the relationship to pelvic floor muscle function in primiparous women at 12 months postpartum: a cross-sectional study', *The Journal of Sexual Medicine*, 12(4), pp. 994-1003.

Test-Club-Bericht (2013) 'Intimina Laselle Vaginalkugeln zur Stärkung der Beckenbodenmuskulatur', *Gesunde Medizin*, (no. 5), pp. 38-39.

Tetzschner, T., Sørensen, M., Lose, G., and Christiansen, J. (1996) 'Pudendal nerve recovery after a non-instrumented vaginal delivery', *International Urogynecology Journal*, 7(2), pp. 102-104.

Thabane, L. and Akhtar-Danesh, N. (2008) 'Guidelines for reporting descriptive statistics in health research', *Nurse Researcher*, 15(2), pp. 72-81.

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L.P., Robson, R., Thabane, M., Giangregorio, L., and Goldsmith, C.H. (2010) 'A tutorial on pilot studies: the what, why and how', *BMC Medical Research Methodology*, 10, p. 1. doi: 10.1186/1471-2288-10-1.

Thabane, L., Mbuagbaw, L., Zhang, S., Samaan, Z., Marcucci, M., Ye, C., Thabane, M., Giangregorio, L., Dennis, B., and Kosa, D. (2013) 'A tutorial on sensitivity analyses in clinical trials: the what, why, when and how', *BMC Medical Research Methodology*, 13, p. 92. doi: 10.1186/1471-2288-13-92.

Thakar, R. and Stanton, S. (2000) 'Regular review: management of urinary incontinence in women', *BMJ*, 321(7272), pp. 1326-1331.

The Analysis Factor (2018) *When does repeated measures ANOVA not work for repeated measures data?* Available at: https://www.theanalysisfactor.com/when-repeated-measures-anova-not-work-for-repeated-measures-data/ (Accessed: 17 July 2018).

The Cochrane Collaboration (2018) *Glossary*. Available at: http://community.cochrane.org/glossary (Accessed: 11 August 2018).

The Cochrane Editorial Resources Committee (2013) 'Data Collection Form (V2)'. Available at: http://training.cochrane.org/authors/resources/checklists-and-forms (Accessed: 8 November 2014).

The Cochrane Pregnancy and Childbirth Group (no date) 'Data Extraction Form'. Available at: http://pregnancy.cochrane.org/author-resources-new-reviews (Accessed: 1 November 2014).

The Joanna Briggs Institute (2011) 'The Joanna Briggs Institute Best Practice Information Sheet: the effectiveness of pelvic floor muscle exercises on urinary incontinence in women following childbirth', *Nursing and Health Sciences*, 13(3), pp. 378-381.

The MathWorks, Inc. (2018) *Generalized linear mixed-effects models*. Available at: https://www.mathworks.com/help/stats/generalized-linear-mixed-effects-models.html (Accessed: 14 July 2018).

The R Foundation (2018) *The R project for statistical computing* [Computer program]. Available at: https://www.r-project.org/ (Downloaded: 29 April 2018).

Theofrastous, J., Wyman, J., Bump, R., McClish, D., Elser, D., Bland, D., and Fantl, J. (2002) 'Effects of pelvic floor muscle training on strength and predictors of response in the treatment of urinary incontinence', *Neurourology and Urodynamics*, 21(5), pp. 486-490.

Thibault-Gagnon, S., Yusuf, S., Langer, S., Wong, V., Sheck, K.L., Martin, A., and Dietz, H.P. (2014) 'Do women notice the impact of childbirth-related levator trauma on pelvic floor and sexual function? Results of an observational ultrasound study', *International Urogynecology Journal*, 25(10), pp. 1389-1398.

Thom, D.H. and Rortveit, G. (2010) 'Prevalence of postpartum urinary incontinence: a systematic review', *Acta Obstetricia et Gynecologica Scandinavica*, 89(12), pp. 1511-1522.

Thoma, A., Farrokhyar, F., McKnight, L., and Bhandari, M. (2010) 'How to optimize patient recruitment', *Canadian Journal of Surgery*, 53(3), pp. 205-210.

Thompson, B. (2002) '"Statistical,""practical," and "clinical": how many kinds of significance do counselors need to consider?', *Journal of Counseling & Development*, 80(1), pp. 64-71.

Thompson, J.A., O'Sullivan, P.B., Briffa, N.K., and Neumann, P. (2006) 'Assessment of voluntary pelvic floor muscle contraction in continent and incontinent women using transperineal ultrasound, manual muscle testing and vaginal squeeze pressure measurements', *International Urogynecology Journal*, 17(6), pp. 624-630.

Thorp, J.M., Norton, P.A., Wall, L.L., Kuller, J.A., Eucker, B., and Wells, E. (1999) 'Urinary incontinence in pregnancy and the puerperium: a prospective study', *American Journal of Obstetrics & Gynecology*, 181(2), pp. 266-273.

Thorpe, K.E., Zwarenstein, M., Oxman, A.D., Treweek, S., Furberg, C.D., Altman, D.G., Tunis, S., Bergel, E., Harvey, I., Magid, D.J., and Chalkidou, K. (2009) 'A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers', *Journal of Clinical Epidemiology*, 62(5), pp. 464-475.

Tickle-Degnen, L. (2013) 'Nuts and bolts of conducting feasibility studies', *American Journal of Occupational Therapy*, 67(2), pp. 171-176.

Tiran, D. (ed.) (2012) *Bailliere's midwives' dictionary e-book*. 12th edn. London: Elsevier Health Sciences UK.

Tong, A., Sainsbury, P., and Craig, J. (2007) 'Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups', *International Journal for Quality in Health Care*, 19(6), pp. 349-357.

Torgerson, D. (2014) Sources of bias in randomised controlled trials [Power Point Presentation]. York: York Trials Unit.

Tortora, G.J. and Derrickson, B.H. (2011) *Principles of anatomy and physiology* (2 vols). 13th edn. Hoboken, NJ: John Wiley & Sons.

Trochim, W.M.K. (2006a) *Convergent & discriminant validity*. Available at: http://www.socialresearchmethods.net/kb/convdisc.htm (Accessed: 16 September 2018).

Trochim, W.M.K. (2006b) *Nonprobability sampling*. Available at: http://www.socialresearchmethods.net/kb/sampnon.php (Accessed: 26 June 2018).

U.S. Department of Health and Human Services (Food and Drug Administration) (2009) *Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims*. Available at: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm193282.pdf (Accessed: 29 June 2018).

U.S. National Library of Medicine (2017) *What are the phases of clinical trials?* Available at: https://www.nih.gov/health-information/nih-clinical-research-trials-you/basics (Accessed: 23 July 2018).

Udier, E. (2014) Skype conversation with Claudia Oblasser. 14 May.

Udier, E. (2018) Skype conversation with Claudia Oblasser. 2 July.

Univalor (no date) *Portable pelvic floor muscle dynamometer*. Available at: http://www.univalor.ca/en/technology/portable-pelvic-floor-muscle-dynamometer (Accessed: 26 June 2018).

USMedia (2017) *Beckenbodentraining mit Hilfsmitteln*. Available at: http://www.rund-ums-baby.de/schwangerschaft/beckenboden-rueckbildung/beckenbodentraining-mit-hilfsmitteln.htm (Accessed: 24 July 2018).

Vaizey, C., Carapeti, E., Cahill, J., and Kamm, M. (1999) 'Prospective comparison of faecal incontinence grading systems', *Gut*, 44(1), pp. 77-80.

'Valsalva maneuver' (2009), in *Mosby's medical dictionary*. 9th edn. Available at: http://medical-dictionary.thefreedictionary.com/Valsalva+maneuver (Accessed: 29 June 2018).

van Brummen, H.J., Bruinse, H.W., van de Pol, G., Heintz, A.P., and van der Vaart, C.H. (2006) 'Defecatory symptoms during and after the first pregnancy: prevalences and associated factors', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 17(3), pp. 224-230.

Van Rooyen, A. (1969) 'Pregnancy and the lower urinary tract: part I-a study of histological changes in the lower urinary tract during pregnancy', *South African Medical Journal*, 43(24), pp. 749-753.

Varney, H., Kriebs, J.M., and Gegor, C.L. (2004) *Varney's midwifery*. 4th edn. Sudbury, MA: Jones and Bartlett Publishers.

Veer West LLC (2018) *FormAssembly*. Available at: https://www.formassembly.com/ (Accessed: 1 June 2018).

Viktrup, L. (2002) 'The risk of lower urinary tract symptoms five years after the first delivery', *Neurourology and Urodynamics*, 21(1), pp. 2-29.

Viktrup, L. and Lose, G. (2000) 'Lower urinary tract symptoms 5 years after the first delivery', *International Urogynecology Journal*, 11(6), pp. 336-340.

Viktrup, L., Lose, G., Rolff, M., and Barfoed, K. (1992) 'The symptom of stress incontinence caused by pregnancy or delivery in primiparas', *Obstetrics & Gynecology*, 79(6), pp. 945-949.

Vodušek, D.B. (2008) 'Neural control of pelvic floor muscles', in Baessler, K., Schüssler, B., Burgio, K.L., Moore, K.H., Norton, P.A. & Stanton, S.L. (eds.) *Pelvic floor re-education: principles and practice.* 2nd edn. London: Springer, pp. 22-35.

Voils, C.I., Levine, E., Gierisch, J.M., Pendergast, J., Hale, S.L., McVay, M.A., Reed, S.D., Yancy, W.S., Bennett, G., Strawbridge, E.M., White, A.C., and Shaw, R.J. (2018) 'Study protocol for Log2Lose: a feasibility randomized controlled trial to evaluate financial incentives for dietary

self-monitoring and interim weight loss in adults with obesity', *Contemporary Clinical Trials*, 65, pp. 116-122.

von Sydow, K. (1999) 'Sexuality during pregnancy and after childbirth: a metacontent analysis of 59 studies', *Journal of Psychosomatic Research*, 47(1), pp. 27-49.

Wai, C.Y., McIntire, D.D., Atnip, S.D., Schaffer, J.I., Bloom, S.L., and Leveno, K.J. (2011) 'Urodynamic indices and pelvic organ prolapse quantification 3 months after vaginal delivery in primiparous women', *International Urogynecology Journal and Pelvic Floor Dysfunction*, 22(10), pp. 1293-1298.

Walker, E. and Nowacki, A.S. (2011) 'Understanding equivalence and noninferiority testing', *Journal of General Internal Medicine*, 26(2), pp. 192-196.

Wallner, C., van Wissen, J., Maas, C.P., Dabhoiwala, N.F., DeRuiter, M.C., and Lamers, W.H. (2008) 'The contribution of the levator ani nerve and the pudendal nerve to the innervation of the levator ani muscles; a study in human fetuses', *European Urology*, 54(5), pp. 1136-1144.

Wason, J.M., Stecher, L., and Mander, A.P. (2014) 'Correcting for multiple-testing in multi-arm trials: is it necessary and is it done?', *Trials*, 15(1), p. 364. doi: 10.1186/1745-6215-15-364.

Weinberg, A. and Chronic Disease Research Center Team (2008) *The EDICT project: policy recommendations to eliminate disparities in clinical trials*. Available at: https://www.researchgate.net/publication/320036453_The_EDICT_Project_Policy_Recommendations_to_Eliminate_Disparities_in_Clinical_Trials (Accessed: 23 July 2018).

Wellkang® Tech Consulting (2017) *What is CE Marking (CE Mark)?* Available at: http://www.ce-marking.org/what-is-ce-marking.html (Accessed: 23 June 2018).

Whitehead, A.L., Julious, S.A., Cooper, C.L., and Campbell, M.J. (2016) 'Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable', *Statistical Methods in Medical Research*, 25(3), pp. 1057-1073.

Whitehead, A.L., Sully, B.G., and Campbell, M.J. (2014) 'Pilot and feasibility studies: is there a difference from each other and from a randomised controlled trial?', *Contemporary Clinical Trials*, 38(1), pp. 130-133.

Wijma, J., Potters, A.E.W., Wolf, B.T., Tinga, D.J., and Aarnoudse, J.G. (2001) 'Anatomical and functional changes in the lower urinary tract during pregnancy', *BJOG: An International Journal of Obstetrics & Gynaecology*, 108(7), pp. 726-732.

Williams, K.S., Assassa, R.P., Gillies, C.L., Abrams, K.R., Turner, D.A., Shaw, C., Haslam, J., Mayne, C., and McGrother, C.W. (2006) 'A randomized controlled trial of the effectiveness of pelvic floor therapies for urodynamic stress and mixed incontinence', *BJU International*, 98(5), pp. 1043-1050.

Williams, M. (2016) 'Guide to ... Feasibility and pilot studies: a guide for NIHR Research Design Service advisors'. Available at: not available any more (Accessed: May 2017).

Williams, M. and Lecouturier, J. (no date) *Pilot and feasibility studies*. Available at: http://www.themedcalls.nihr.ac.uk/__data/assets/pdf_file/0019/114175/LTC-CYP-Pilot-and-Feasibility-studies.pdf (Accessed: 8 January 2015).

Wilmore, J.H. and Costill, D. (2004) *Physiology of sport and exercise*. 3rd edn. Champaign, IL: Human Kinetics.

Wilson, P.D. and Herbison, G.P. (1998) 'A randomized controlled trial of pelvic floor muscle exercises to treat postnatal urinary incontinence', *International Urogynecology Journal*, 9(5), pp. 257-264.

Wimmer-Puchinger, B., Bässler, D.C., Beurle, M.A., and Raunig, M.J. (2013) *Psychosoziale Einflussfaktoren auf Geburtsmethoden und Zufriedenheit: Eine multizentrische empirische Studie an Frauen im Wochenbett*. Available at: https://www.wien.gv.at/gesundheit/beratung-vorsorge/frauen/frauengesundheit/pdf/psychoziale-einflussfaktoren-geburtsmethoden.pdf (Accessed: 29 June 2018).

Wisniewski, P.M. and Wilkinson, E.J. (1991) 'Postpartum vaginal atrophy', *American Journal of Obstetrics & Gynecology*, 165(4), pp. 1249-1254.

Woodley, S.J., Boyle, R., Cody, J.D., Mørkved, S., and Hay-Smith, E.J.C. (2017) 'Pelvic floor muscle training for prevention and treatment of urinary and faecal incontinence in antenatal and postnatal women', *Cochrane Database of Systematic Reviews*, (12), p. CD007471.

World Health Organization (2010) *WHO technical consultation on postpartum and postnatal care*. Available at: http://apps.who.int/iris/handle/10665/70432 (Accessed: 26 June 2018).

World Health Organization (2014) *WHO recommendations on postnatal care of the mother and newborn*. Available at: http://apps.who.int/iris/bitstream/10665/97603/1/9789241506649_eng.pdf (Accessed: 29 June 2018).

World Health Organization (2018a) *BMI classification*. Available at: http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi (Accessed: 23 July 2018).

World Health Organization (2018b) *Medical device – full definition*. Available at: http://www.who.int/medical_devices/full_deffinition/en/ (Accessed: 26 June 2018).

World Medical Association (2018) *World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects*. Available at: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ (Accessed: 26 June 2018).

Yeniel, A. and Petri, E. (2014) 'Pregnancy, childbirth, and sexual function: perceptions and facts', *International Urogynecology Journal*, 25(1), pp. 5-14.

ZHAW Zürcher Hochschule für Angewandte Wissenschaften (2016) E-mail to Claudia Oblasser. 30 May.

ZHAW Zürcher Hochschule für Angewandte Wissenschaften (no date) *Machbarkeitsstudie für die Entwicklung einer Einlegesohle für den kyBoot mit sensomotorischen Elementen zur Aktivierung des Beckenbodens nach der Geburt*. Available at: https://www.zhaw.ch/no_cache/de/forschung/forschungsdatenbank/projektdetail/projektid/867/ (Accessed: 21 June 2018).

Ziman, J. (2000) *Real science: what it is and what it means*. Cambridge: Cambridge University Press.

Zizzi, P.T., Trevisan, K.F., Leister, N., Cruz, C.d.S., and Riesco, M.L.G. (2017) 'Women's pelvic floor muscle strength and urinary and anal incontinence after childbirth: a cross-sectional study', *Revista da Escola de Enfermagem da USP*, 51, p. e03214. doi: 10.1590/s1980-220x2016209903214.