



## City Research Online

### City, University of London Institutional Repository

---

**Citation:** Kather, J. N., Berghoff, A. S., Ferber, D., Suarez-Carmona, M., Reyes-Aldasoro, C. C., Valous, N. A., Rojas-Moraleda, R., Jäger, D. & Halama, N. (2018). Large-scale database mining reveals hidden trends and future directions for cancer immunotherapy. *Oncolimmunology*, 7(7), e1444412. doi: 10.1080/2162402x.2018.1444412

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/19561/>

**Link to published version:** <https://doi.org/10.1080/2162402x.2018.1444412>

**Copyright:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

---

City Research Online:

<http://openaccess.city.ac.uk/>

[publications@city.ac.uk](mailto:publications@city.ac.uk)

---

# Large-scale database mining reveals hidden trends and future directions for cancer immunotherapy

Jakob Nikolas Kather<sup>1,2,3</sup>, Anna Sophie Berghoff<sup>1,4</sup>, Dyke Ferber<sup>1,3</sup>, Meggy Suarez-Carmona<sup>1,3</sup>,  
Constantino Carlos Reyes-Aldasoro<sup>5</sup>, Nektarios A. Valous<sup>1,3</sup>, Rodrigo Rojas-Moraleda<sup>1,3</sup>,  
Dirk Jäger<sup>1,2,3</sup>, Niels Halama<sup>1,2,3</sup>

<sup>1</sup> Department of Medical Oncology and Internal Medicine VI, National Center for Tumor Diseases, University Hospital Heidelberg, Heidelberg, Germany

<sup>2</sup> German Cancer Consortium (DKTK), Heidelberg, Germany

<sup>3</sup> Clinical Cooperation Unit Applied Tumor Immunity, German Cancer Research Center (DKFZ), Heidelberg, Germany

<sup>4</sup> Clinical Unit for Experimental Oncology Therapy, Thoraxklinik, University of Heidelberg, Heidelberg, Germany

<sup>5</sup> School of Mathematics, Computer Science and Engineering, City University of London, London, UK

Corresponding author information:

Niels Halama, MD

Medical Oncology, National Center for Tumor Diseases (NCT)

Im Neuenheimer Feld 460; 69120 Heidelberg, Germany

niels.halama@nct-heidelberg.de, +4962215638397

1 **Abstract**

2 Cancer immunotherapy has fundamentally changed the landscape of oncology in recent years  
3 and significant resources are invested into immunotherapy research. It is in the interests of re-  
4 searchers and clinicians to identify promising and less promising trends in this field in order to  
5 rationally allocate resources. This requires a quantitative large-scale analysis of cancer immuno-  
6 therapy related databases.

7 We developed a novel tool for text mining, statistical analysis and data visualization of scientific  
8 literature data. We used this tool to analyze 72002 cancer immunotherapy publications and  
9 1469 clinical trials from public databases. All source codes are available under an open access  
10 license.

11 The contribution of specific topics within the cancer immunotherapy field has markedly shifted  
12 over the years. We show that the focus is moving from cell-based therapy and vaccination to-  
13 wards checkpoint inhibitors, with these trends reaching statistical significance. Rapidly growing  
14 subfields include the combination of chemotherapy with checkpoint blockade. Translational  
15 studies have shifted from hematological and skin neoplasms to gastrointestinal and lung cancer  
16 and from tumor antigens and angiogenesis to tumor stroma and apoptosis.

17 This work highlights the importance of unbiased large-scale database mining to assess trends in  
18 cancer research and cancer immunotherapy in particular. Researchers, clinicians and funding  
19 agencies should be aware of quantitative trends in the immunotherapy field, allocate resources  
20 to the most promising areas and find new approaches for currently immature topics.

21

22 **Keywords:** Cancer immunotherapy, database mining, checkpoint inhibition, gastrointestinal  
23 cancer, lung cancer, translational research

## 24 Introduction

25 Cancer immunotherapy is widely regarded as one of the most promising approaches for treat-  
26 ing metastatic cancer.<sup>1</sup> It has been in the focus of basic, translational and clinical research for  
27 years and significant resources have been invested in finding new immunotherapy treatments  
28 with clinical efficacy.

29 Anecdotally, most clinicians and researchers in the field are aware that clinical translation has  
30 not been equally successful for each subfield over the last years. For example, it is well-known  
31 that therapeutic vaccines were intensely investigated and shaped immunotherapy for years but  
32 have not yet made a direct clinical impact. Also, immunotherapy quickly reached clinical appli-  
33 cation in melanoma<sup>2</sup>, while gastrointestinal cancer types are still lagging behind.<sup>3</sup> These shifts  
34 within the cancer immunotherapy field are highly relevant for clinicians, researchers and fund-  
35 ing agencies. However, until now, these changes have not been quantified in a way that allows  
36 an unbiased assessment of past and possible future trends.

37 In the present study, we quantified the development of the cancer immunotherapy field from  
38 1986 to 2017 to reveal previously hidden trends. This type of quantitative and unbiased analysis  
39 is of high interest to researchers and clinicians because it can guide the allocation of resources  
40 for future research and clinical trials. Specifically, we focused on the comparison of treatment  
41 approaches, translational research topics and different tumor entities (organ of the primary  
42 tumor, according to the International Statistical Classification of Diseases and Related Health  
43 Problems, ICD-10). Among various types of cancer immunotherapy<sup>4</sup>, we looked at the devel-  
44 opment of oncolytic viruses<sup>5</sup>, cell-based therapies<sup>6</sup>, therapeutic vaccines<sup>7</sup>, checkpoint inhibi-  
45 tors<sup>8, 9</sup> as well as chemotherapy and radiation therapy. These treatment types were separately  
46 analyzed for all tumor entities in order find out which approaches would be most promising in  
47 specific entities in the future. To quantify developments in basic and translational cancer re-  
48 search, we included a wide range of topics such as the combination of immunotherapy with  
49 stroma<sup>10</sup> and cancer-associated fibroblasts<sup>11</sup>, angiogenesis<sup>12</sup>, tumor-specific antigens<sup>13</sup>, neoan-  
50 tigen<sup>14</sup>, microbiota<sup>15</sup>, drug resistance<sup>16</sup>, myeloid cells<sup>17</sup>, stem cells<sup>18</sup>, epigenetics<sup>19</sup>, cell death  
51 and autophagy<sup>20, 21</sup> as well as metabolism<sup>22</sup>. All trends were analyzed over time, keeping in

52 mind that the field was profoundly changed by landmark events such as the first clinical report  
53 of effective checkpoint inhibition in cancer patients in 2003.<sup>23, 24</sup> Inhibitors of immune receptors  
54 and ligands are currently the largest class of approved immunotherapy drugs.<sup>25, 26</sup> To investi-  
55 gate this subfield in detail, we used a graph-based approach to visualize which of these check-  
56 point pathways was in the focus of research efforts during the last years. Also, this analysis was  
57 used to identify promising combination approaches to target checkpoint signaling pathways.

58 In short, we present a novel method for data collection, analysis and visualization of changing  
59 trends in cancer immunotherapy from 1986 to 2017 and discuss their implications.

## 60 **Methods**

### 61 Database queries

62 Based on previous literature reviews and other publicly available resources, we manually curat-  
63 ed a list of keywords to enable the comparison of different tumor entities (organ of the primary  
64 tumor, e.g. brain, breast, sarcoma, etc., complete list in Suppl. Table 1), treatment approaches  
65 (e.g. adoptive cell transfer, oncolytic viruses, checkpoint inhibition, etc., complete list in Suppl.  
66 Table 2), translational research topics (e.g. apoptosis, stem cells, epigenetics, etc., complete list  
67 in Suppl. Table 3) and cell types (e.g. myeloid, lymphoid, etc., complete list in Suppl. Table 4).  
68 Resources for therapeutic agents were the “NIH: A to Z List of Cancer Drugs” (retrieved from  
69 <https://www.cancer.gov/about-cancer/treatment/drugs> on 11 Nov 2017) and all FDA approvals  
70 2016 and 2017 (retrieved from  
71 <https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm279174.htm> on 11 Nov  
72 2017). Publication data were automatically mined from MEDLINE, the database of the United  
73 States National Library of Medicine (NLM), and its related search engine PubMed  
74 (<https://pubmed.gov>). Furthermore, we analyzed all cancer immunotherapy clinical trials regis-  
75 tered in the official US (<https://clinicaltrials.gov>) database. PubMed articles were identified by  
76 the following master search keyword: ("tumor"[All Fields] OR "tumors"[All Fields] OR "neo-  
77 plasms"[MeSH Terms] OR "neoplasms"[All Fields] OR "cancer"[All Fields]) AND ("immunothera-  
78 py"[MeSH Terms] OR "immunotherapy"[All Fields]), in a similar way to a previously published

79 study.<sup>27</sup> For clinical trials, the master keyword was: “cancer immunotherapy”. For clinical trials,  
80 all accessible trial metadata (title, description and structured information) was downloaded  
81 from respective databases. All database queries were made in November 2017.

## 82 Data analysis

83 All data analyzes and visualizations were conducted with self-developed MATLAB scripts  
84 (R2017a, MathWorks, Natick, MA, USA). Data were normalized to the number of total immuno-  
85 therapy articles (or trials, respectively) in each year. Data points were smoothed with a moving  
86 average filter (lowpass filter with a coefficient equal to the reciprocal of the time span and a  
87 window size of five years). All scripts are released open-source and are available in: [DOI will be  
88 inserted after acceptance].

## 89 Trumpet plot

90 To illustrate the temporal variation of the incidence of keyword groups, we used the self-  
91 developed “*trumpet plot*”. Normalized and smoothed timelines were visualized as the height of  
92 a “trumpet” shape in a 2D. In 3D, the diameter of a cylinder represented the normalized num-  
93 ber of research items in a given year with time as the vertical axis. Perceptually optimized col-  
94 our scales from the “Color Brewer” project were used to visualize data.<sup>28</sup>

## 95 Graph-based analysis and network plot

96 To investigate the degree of connectivity between similar keywords in a specific subfield, we  
97 used a graph-based analysis. This was employed for keywords that represented different im-  
98 mune checkpoint molecules e.g. PD-1, PD-L1, CTLA-4, CD80, etc. (full list in Suppl. Table 1). Each  
99 keyword was represented by a node which was visualized as a circle. The size and color of the  
100 circle depicted the number of research items matching this keyword. The distance between the  
101 node and the width of the connecting edge represented the co-occurrence of two keywords.  
102 Logarithmic scaling was used for the circle size and the edge width. Isolated nodes without any  
103 connection to other nodes were discarded. Low-abundant nodes (< 10 hits) and edges were al-  
104 so discarded.

## 105 **Results**

### 106 Shift from vaccination to checkpoint inhibition in clinical and translational studies

107 First, we analyzed the contribution of major treatment types to the cancer immunotherapy lit-  
108 erature. In the PubMed database, chemotherapy was the most frequent treatment that articles  
109 could be matched to (33% in 2017, Figure 1A). Checkpoint inhibition grew significantly (indicat-  
110 ed by a + in the graphs) from 2015 and was the second most abundant treatment type in 2017.  
111 Therapeutic vaccination as a form of cancer immunotherapy dropped from position 1 to posi-  
112 tion 3 in 2017, with significant decrease (indicated by a diamond in the graphs) between 2015  
113 and 2017. These trends were even more pronounced in clinical trials where checkpoint inhibi-  
114 tion was matched in more than 50% of all items in 2017, chemotherapy being second with 26%  
115 and vaccination steadily dropping to only 9% of clinical trials in 2017 (Figure 2A). Adoptive cell-  
116 based therapies (including chimeric antigen receptor [CAR] T-cells) contributed to 15% of all  
117 research items in 2017 and to 7% of all clinical trials (Figure 1A and Figure 2A).

### 118 Lung and gastrointestinal cancer as prime targets for immunotherapy

119 Next, we analyzed cancer immunotherapy research efforts for each tumor entity. In articles in-  
120 dexed in PubMed, hematological neoplasias (hema.) were the prime immunotherapy target un-  
121 til 2015/2016, but has decreased significantly since, yielding to skin and gastrointestinal (GI)  
122 neoplasms (Figure 1B). Among the top five tumor entities (skin, GI, hema., respiratory-thoracic  
123 [lung] and urinary tract), only lung and GI showed a significant growth in the last five years  
124 (Figure 1B). This pattern matched clinical trial data (Figure 2B) where lung and GI tumors were  
125 the top two cancer entities by far. Again, hematological neoplasms rapidly (and in one year sig-  
126 nificantly) decreased in importance; also, sarcoma continuously decreased in importance over  
127 the years (Figure 2B).

128 Subsequently, we asked how the different therapy approaches were reflected in each major  
129 tumor entity. In the research literature, checkpoint inhibitors have increased in importance in  
130 the last five years in all top five tumor entities (Figure 1D). The reverse trend can be observed in  
131 vaccination and chemotherapy, although these still have a large presence. Much more pro-

132 nounced effects were observed in clinical trials (Figure 2D): Here, lung and GI neoplasms were  
133 the two most dynamically growing field with growth in skin cancer reaching a plateau and he-  
134 matological neoplasms vanishing almost completely.

#### 135 A transient 1990s interest in myeloid cells left no trace in the clinic

136 Cancer immunotherapy aims to (re)invigorate the host immune response against malignant  
137 cells and all types of cancer immunotherapy use cells in the tumor microenvironment as their  
138 effectors. We analyzed the quantitative contribution of cell types in the immunotherapy litera-  
139 ture. Items related to myeloid cells significantly increased its presence in PubMed in the late  
140 1990s (Figure 1C), matching a large contribution to clinical trials at that time (Figure 2C). How-  
141 ever, this transient interest in myeloid cells plateaued in the scientific literature and rapidly de-  
142 creased in clinical trials. Not surprisingly, lymphoid cells were the largest single group of cells in  
143 2017 in scientific publications and clinical trials.

#### 144 Revival of radiation and chemo-immunotherapy

145 Having analyzed major trends among treatment types, cancer types and cell types, we looked  
146 for non-obvious trends in the dataset. We found that among treatment types, radiation was  
147 only at position five in scientific articles (Figure 1A) but at position three in clinical trials (Figure  
148 2A). In both cases, the growth rate in 2017 significantly exceeded that of previous years. These  
149 trends followed a decrease during the early 2000s in radiation therapy in articles and clinical  
150 trials (Figure 1A and Figure 2A). Based on these data, we conclude that we are currently wit-  
151 nessing a revival of the use of radiation in cancer immunotherapy.

152 We hypothesized that other non-obvious trends might be hidden in treatment combinations  
153 and therefore analyzed co-occurrence of treatment types in clinical trials (Figure 3A). In this  
154 analysis, the diagonal of the matrix corresponds to Figure 2A. We found that the only markedly  
155 increasing treatment combination is chemotherapy plus checkpoint inhibition (Figure 3A). In  
156 contrast, virtually no registered clinical trials investigate the combinations vaccination plus  
157 checkpoint inhibition or adoptive cellular therapy plus checkpoint inhibition.

158 Stroma and apoptosis in gastrointestinal cancer

159 Our automatic approach for database mining allowed for an analysis of translational research  
160 topics per tumor type. For clarity, only a part of this analysis is shown in Figure 3B. We found  
161 that among translational research topics in immunotherapy articles, angiogenesis is decreasing  
162 in importance in all major cancer entities. In contrast, apoptosis (and other forms of cell death  
163 as well as autophagy) is rapidly gaining ground in GI, lung and skin cancer (Figure 3B). Interest-  
164 ingly, the quantitative contribution of cancer stroma to immunotherapy articles is stagnating or  
165 decreasing in all major cancer entities except GI cancer (Figure 3B). Complementing our above-  
166 described finding that GI cancer is one of the most dynamically growing research topics in im-  
167 munotherapy, we conclude that especially apoptosis and stroma are promising subfields in this  
168 entity.

169 Translational activities vary considerably between tumor types

170 Our next step was to examine the following question: how were preclinical research efforts,  
171 measured by the number of indexed items on PubMed, translated into clinical trials? To give a  
172 specific answer for all therapy types and major cancer entities, we compared timelines for mul-  
173 tiple keywords in PubMed and clinical trial databases. We analyzed the number of clinical trials  
174 in the last five years (2012-2016) and normalized these numbers to the respective number of  
175 PubMed research items in the preceding five years. Among all therapy types, immune check-  
176 point inhibition stood out in terms of translational efficiency with close to 0.2 clinical trials per  
177 research paper in the reference periods (Figure 4A). Looking at various tumor entities, the dif-  
178 ferences in translational efficiency were not as large (Figure 4B). Highest translational efficiency  
179 was visible in immunotherapy of gastrointestinal and respiratory neoplasms while a low transla-  
180 tional efficiency was seen in hematological malignancies with just 0.02 clinical trials per article  
181 (Figure 4B).

182 Another way of comparing the translational efficiency of immunotherapy subfields is to look at  
183 the development of clinical phase 1/2/3 trials over time. We matched all cancer immunothera-  
184 py trials registered at clinicaltrials.gov and all PubMed articles (when applicable) to one or more  
185 clinical phases. In the timelines in Figure 5A, a small and stable percentage of PubMed articles

186 can be matched to any clinical trial phase over time. Within registered clinical trials (Figure 5B),  
187 phase 1 and 2 trials are slowly increasing with phase 3 trials decreasing at the same time. How-  
188 ever, in general, no pronounced trends were visible in this analysis. This picture changed mark-  
189 edly when analyzing clinical trials for each major tumor entity (Figure 5C): Phase 1 and 2 trials  
190 were rapidly increasing in gastrointestinal and lung cancer in the last five to ten years, but not  
191 in other major tumor entities. These data match our above-mentioned finding that GI and lung  
192 cancer are the most translationally active fields as compared to skin cancer, hematological neo-  
193 plasias and other major cancer types.

#### 194 Immune-checkpoint networks

195 Based on above-described results we concluded that checkpoint inhibition makes the largest  
196 quantitative contribution to research papers and clinical trials in immunotherapy research and  
197 is also the most efficient subfield in terms of clinical translation. Therefore, we performed a  
198 more specific analysis and asked how the contribution and intertwining of immune checkpoint  
199 molecules and drugs developed over time. Based on our timeline analysis (Figure 1A) we esti-  
200 mated that around 2011, the increase in checkpoint inhibition publications started. We there-  
201 fore used the following time frames, 1986-2010 and 2011-2016, to compare co-occurrence of  
202 checkpoint molecules in PubMed articles. These comparisons are shown in Figure 6 as network  
203 plots. In 1986 to 2010, CD80 had the highest prevalence (Figure 6A) and a cluster around  
204 CD80/CD86/CD28/CD40L/CD40 dominated the immune checkpoint landscape in PubMed arti-  
205 cles. In 2011 to 2016, a marked change was evident and PD-1/PD-L1, which were previously in  
206 the periphery of the network, and CTLA-4, which remained in the center, made by far the larg-  
207 est contribution (Figure 6B). Interestingly, CD80 (B7-1) still occupied a central “hub” position,  
208 linking two distant parts of the network with each other.

## 209 **Discussion**

210 Tumor immunotherapy research is a dynamically evolving field and has undergone profound  
211 changes in the last three decades. While these developments might be implicitly known by re-  
212 searchers who have been deeply involved in the field for a long time, they are probably not ap-

213 parent to most clinicians and scientists who are now confronted with immunotherapy. Moreo-  
214 ver, researchers and clinicians working in the field may have cognitive biases and therefore may  
215 not be aware of well and poorly performing subfields of immunotherapy research. In this paper,  
216 we presented a quantitative, objective and comprehensive analysis of the changes in tumor  
217 immunotherapy research over time which can serve as a rational basis for further discussions.

218 Skin cancer (mainly melanoma) was the first tumor entity to have effective immunotherapy  
219 agents approved and is still in the focus of research papers. Yet, clinical trials now focus on gas-  
220 trointestinal and respiratory cancers, two major disease classes associated with significant  
221 morbidity and mortality. Translational research means that new knowledge should be effective-  
222 ly transferred to the clinic.<sup>29</sup> Researchers pursuing translational research will therefore meet  
223 this aim more easily in an area where translation has been shown to be feasible. By extrapolat-  
224 ing these current trends, translational research efforts would be most fruitful in gastrointestinal  
225 and respiratory cancer.

226 As a word of caution, we should also acknowledge that many unexpected breakthroughs come  
227 from previously unnoticed areas in biomedical research. Also, not all ongoing research efforts  
228 might be reflected by PubMed publications or registered clinical trials. Yet, for the tedious pro-  
229 cess of using research results from the laboratory to improve treatments in the clinic, a struc-  
230 tured and objective projection of future trends can be very useful. Our data-driven analytics  
231 approach provides a starting basis for such endeavors .

## 232 **Acknowledgements**

233 The authors are very grateful to Dr. Charles Neu for proofreading the article.

## 234 **References**

- 235 1. Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. *J Clin Invest*  
236 2015; 125:3335-7.
- 237 2. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival  
238 with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010; 363:711-23.
- 239 3. Jager D, Halama N, Zornig I, Klug P, Krauss J, Haag GM. Immunotherapy of Colorectal Cancer.  
240 *Oncol Res Treat* 2016; 39:346-50.
- 241 4. Galluzzi L, Vacchelli E, Bravo-San Pedro JM, Buque A, Senovilla L, Baracco EE, et al. Classification  
242 of current anticancer immunotherapies. *Oncotarget* 2014; 5:12472-508.
- 243 5. Lawler SE, Speranza MC, Cho CF, ChioCCA EA. Oncolytic Viruses in Cancer Treatment: A Review.  
244 *JAMA Oncol* 2016.
- 245 6. Sadelain M, Riviere I, Brentjens R. Targeting tumours with genetically enhanced T lymphocytes.  
246 *Nat Rev Cancer* 2003; 3:35-45.
- 247 7. Ribas A, Butterfield LH, Glaspy JA, Economou JS. Current developments in cancer vaccines and  
248 cellular immunotherapy. *J Clin Oncol* 2003; 21:2415-32.
- 249 8. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*  
250 2012; 12:252-64.
- 251 9. Postow MA, Callahan MK, Wolchok JD. Immune Checkpoint Blockade in Cancer Therapy. *J Clin*  
252 *Oncol* 2015; 33:1974-82.
- 253 10. Kammertoens T, Schuler T, Blankenstein T. Immunotherapy: target the stroma to hit the tumor.  
254 *Trends Mol Med* 2005; 11:225-31.
- 255 11. Busch S, Andersson D, Bom E, Walsh C, Stahlberg A, Landberg G. Cellular organization and  
256 molecular differentiation model of breast cancer-associated fibroblasts. *Mol Cancer* 2017; 16:73.
- 257 12. Villanueva MT. Immunotherapy: T cells get a ride. *Nat Rev Cancer* 2017; 17:333.
- 258 13. Rosenberg SA. A new era for cancer immunotherapy based on the genes that encode cancer  
259 antigens. *Immunity* 1999; 10:281-7.
- 260 14. Yarchoan M, Johnson BA, 3rd, Lutz ER, Laheru DA, Jaffee EM. Targeting neoantigens to augment  
261 antitumour immunity. *Nat Rev Cancer* 2017; 17:209-22.
- 262 15. Roy S, Trinchieri G. Microbiota: a key orchestrator of cancer therapy. *Nat Rev Cancer* 2017;  
263 17:271-85.
- 264 16. Restifo NP, Smyth MJ, Snyder A. Acquired resistance to immunotherapy and future challenges.  
265 *Nat Rev Cancer* 2016; 16:121-6.
- 266 17. Engblom C, Pfirschke C, Pittet MJ. The role of myeloid cells in cancer therapies. *Nat Rev Cancer*  
267 2016; 16:447-62.
- 268 18. Gammaitoni L, Leuci V, Mesiano G, Giraudo L, Todorovic M, Carnevale-Schianca F, et al.  
269 Immunotherapy of cancer stem cells in solid tumors: initial findings and future prospective. *Expert Opin*  
270 *Biol Ther* 2014; 14:1259-70.
- 271 19. Prachayasittikul V, Prathipati P, Pratiwi R, Phanus-Umporn C, Malik AA, Schaduangrat N, et al.  
272 Exploring the epigenetic drug discovery landscape. *Expert Opin Drug Discov* 2017; 12:345-62.

- 273 20. Fulda S, Vucic D. Targeting IAP proteins for therapeutic intervention in cancer. *Nat Rev Drug*  
274 *Discov* 2012; 11:109-24.
- 275 21. Pan H, Chen L, Xu Y, Han W, Lou F, Fei W, et al. Autophagy-associated immune responses and  
276 cancer immunotherapy. *Oncotarget* 2016; 7:21235-46.
- 277 22. Cairns RA, Mak TW. The current state of cancer metabolism. *Nat Rev Cancer* 2016; 16:613-4.
- 278 23. Phan GQ, Yang JC, Sherry RM, Hwu P, Topalian SL, Schwartzentruber DJ, et al. Cancer regression  
279 and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with  
280 metastatic melanoma. *Proc Natl Acad Sci U S A* 2003; 100:8372-7.
- 281 24. Hodi FS, Mihm MC, Soiffer RJ, Haluska FG, Butler M, Seiden MV, et al. Biologic activity of  
282 cytotoxic T lymphocyte-associated antigen 4 antibody blockade in previously vaccinated metastatic  
283 melanoma and ovarian carcinoma patients. *Proc Natl Acad Sci U S A* 2003; 100:4712-7.
- 284 25. Dong H, Strome SE, Salomao DR, Tamura H, Hirano F, Flies DB, et al. Tumor-associated B7-H1  
285 promotes T-cell apoptosis: a potential mechanism of immune evasion. *Nat Med* 2002; 8:793-800.
- 286 26. Brahmer JR, Drake CG, Wollner I, Powderly JD, Picus J, Sharfman WH, et al. Phase I study of  
287 single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity,  
288 pharmacodynamics, and immunologic correlates. *J Clin Oncol* 2010; 28:3167-75.
- 289 27. Reyes-Aldasoro CC. The proportion of cancer-related entries in PubMed has increased  
290 considerably; is cancer truly "The Emperor of All Maladies"? *PLoS One* 2017; 12:e0173671.
- 291 28. Harrower M, Brewer CA. ColorBrewer.org: an online tool for selecting colour schemes for maps.  
292 *Cartogr J* 2003; 40:27-37.
- 293 29. Woolf SH. The meaning of translational research and why it matters. *JAMA* 2008; 299:211-3.

294

295 **Figure Legends**

296 **Figure 1: Trends in PubMed publications from 1986 to 2017 by topic.** This figure summarizes all  
297 PubMed listed cancer immunotherapy articles grouped by category. (A) Among all cancer im-  
298 munotherapy articles published in 2017, 33% referred to one or more specific chemotherapy  
299 drugs (bottom shape). This proportion was roughly constant over three decades. Contrariwise,  
300 checkpoint inhibition was almost absent before 2010, showing an accelerating growth after-  
301 wards. (B) Hematological neoplasms were the most commonly investigated immunotherapy  
302 target until 2014, when they were overtaken by gastrointestinal and skin neoplasms. (C) Among  
303 all major cell types in the tumor microenvironment, myeloid cells were rapidly gaining interest  
304 around the year 2000. Afterwards, no significant change whatsoever was observed. (D) Bivari-  
305 ate analysis of treatment types versus cancer types in PubMed cancer immunotherapy publica-  
306 tions. Checkpoint inhibition shows a markedly increasing trend (“trumpet”) in skin, respiratory,  
307 but also urinary tract and gastrointestinal cancer. (A-C) + significant rise of growth rate within  
308 one year (anomaly >95%),  $\diamond$  significant decrease of one-year growth rate (anomaly < 5%).

309

310 **Figure 2: Trends in clinical trials from 2000 to 2017 per topic.** This figure summarizes all regis-  
311 tered clinical trials of cancer immunotherapy grouped by category. (A) Immune checkpoint in-  
312 hibition has rapidly become the most common therapy approach between 2010 and 2017. At  
313 the same time, vaccination approaches have greatly diminished, being subject to only 9% of  
314 clinical immunotherapy trials in 2017. (B) As in PubMed publications, hematological neoplasms  
315 have markedly lost ground, yielding to gastrointestinal and respiratory neoplasms in recent  
316 years. (C) Among all major cell types in the tumor microenvironment, myeloid cells were in the  
317 focus of research interest around 2000, diminishing afterwards and only being investigated in  
318 2% of immunotherapy clinical trials in 2017. (D) Bivariate plot of treatment types versus cancer  
319 types in cancer immunotherapy clinical trials. Checkpoint inhibition shows an increasing trend  
320 (“trumpet”) in respiratory and gastrointestinal cancer. (A-C) + significant rise of growth rate  
321 within one year (anomaly >95%),  $\diamond$  significant decrease of one-year growth rate (anomaly <  
322 5%).

323

324 **Figure 3: Emerging immunotherapy paradigms.** (A) Co-occurrence of cancer immunotherapy  
325 treatment approaches in clinical trials between 2000 and 2017. On the diagonal, the develop-  
326 ment of individual treatment approaches is shown with checkpoint inhibition displaying a rapid  
327 increase. Off the diagonal, treatment combinations are shown with chemotherapy and check-  
328 point inhibition being the most common and rapidly growing combination. (B) This bivariate  
329 plot shows cancer immunotherapy trends grouped by translational research topics and major  
330 cancer types based on all PubMed publications between 1986 and 2017. Among signalling,  
331 stroma, apoptosis and angiogenesis, apoptosis is the most rapidly growing topic in all major  
332 cancer entities except hematological neoplasms. Stroma and signalling are most rapidly increas-  
333 ing in gastrointestinal cancer.

334

335 **Figure 4: Translational efficiency.** We asked how the number of research publications influ-  
336 ences the number of clinical trials in subsequent years. To this end, we analyzed PubMed arti-  
337 cles for specific fields in a five-year period (2006-2011) and evaluated the number of matching  
338 US-registered clinical trials in the following five years (2012-2016). This yields a measure of  
339 translational efficiency (clinical trials per research publication). (A) Among therapy types, im-  
340 mune checkpoint inhibitors had the highest translational efficiency with approximately 0.2 trials  
341 per publication. Scientific findings in vaccination and cell-based therapy were not efficiently  
342 translated to the clinic. (B) Among major tumor entities, translational efficiency was highest for  
343 gastrointestinal tumors and lowest for hematological and lymphoid malignancies (hema.). It is  
344 in the interest of the research community to increase translational efficiency in these low-  
345 performing fields.

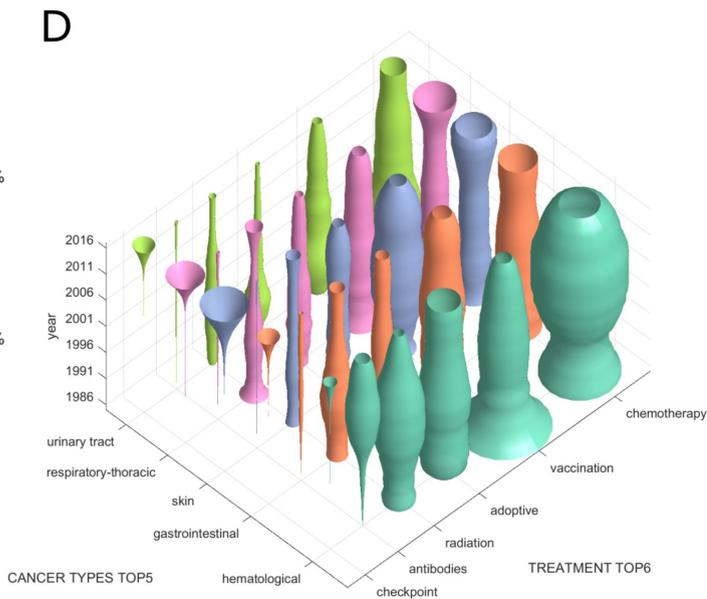
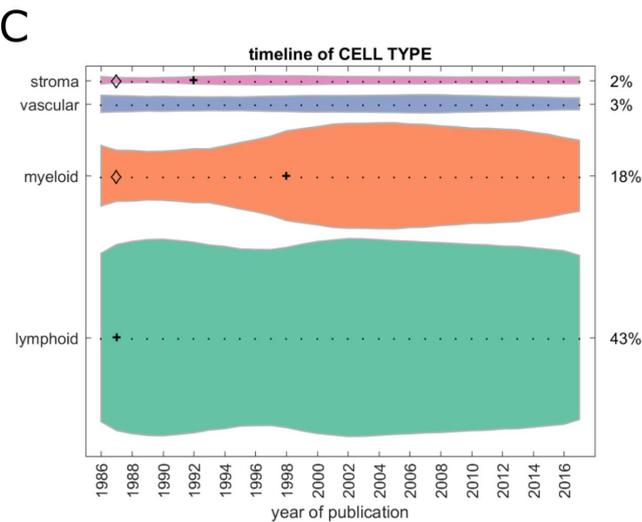
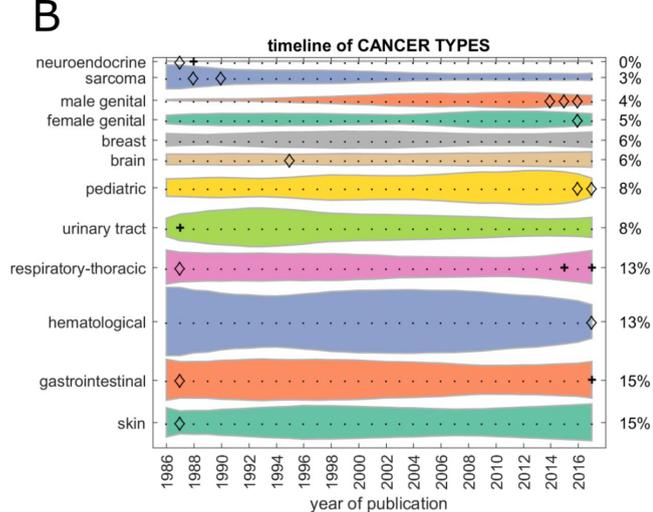
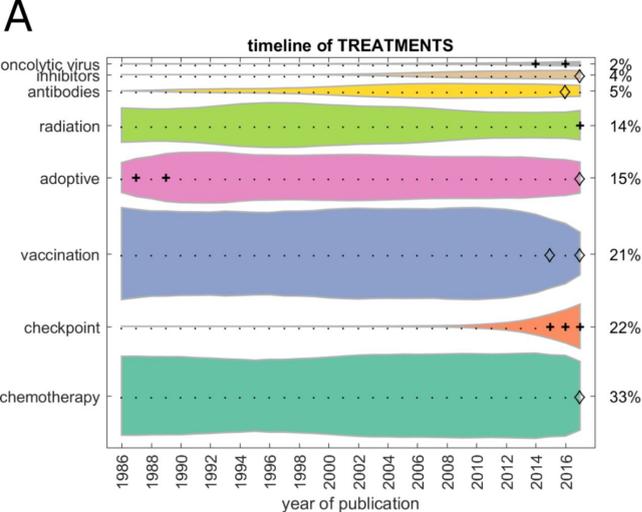
346

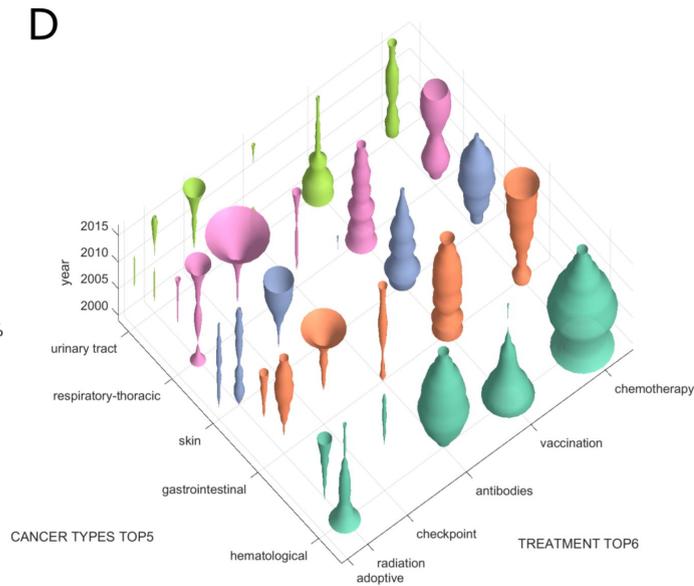
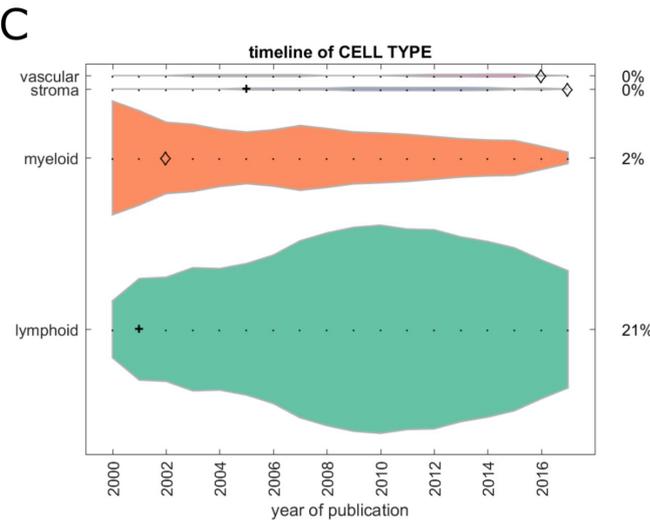
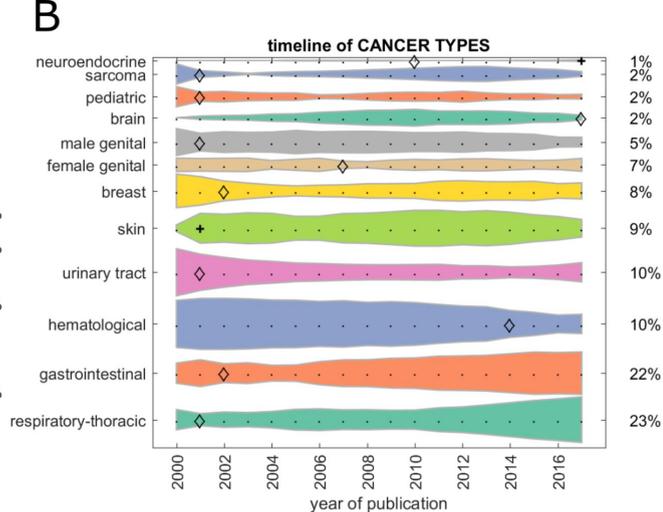
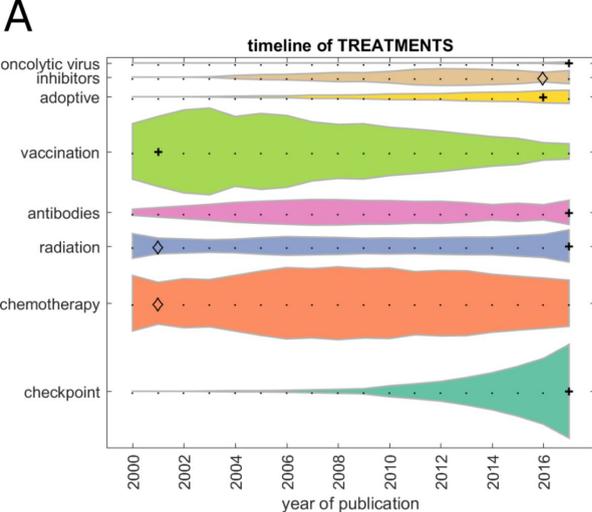
347 **Figure 5: Clinical trial phases.** This figure shows the development of clinical trials in phase  
348 1/2/3 over time. (A) PubMed articles matching any clinical phase. Only a fraction of PubMed  
349 listed articles can be matched to a clinical phase and the proportions between the phases have  
350 not changed significantly in the last 20 years. (B) Clinical trials matching any clinical phase, or-  
351 dered by group size, from bottom to top: phase 2, phase 1, phase 3. Some trials could be  
352 matched to multiple phases so that the percentages in 2017 do not necessarily add up to 100%.  
353 Phase 2 trials are most abundant and phase 1 trials are slowly growing, albeit not significantly.  
354 (C) In stark contrast to the slow overall growth dynamic of clinical trials in the above panels, this  
355 panel shows marked changes in clinical trials per cancer entity over time. In gastrointestinal  
356 cancer and respiratory-thoracic cancers, phase 1 and 2 trials are currently showing pronounced  
357 increase. (A+B) + significant rise of growth rate within one year (anomaly >95%),  $\diamond$  significant  
358 decrease of one-year growth rate (anomaly < 5%).

359

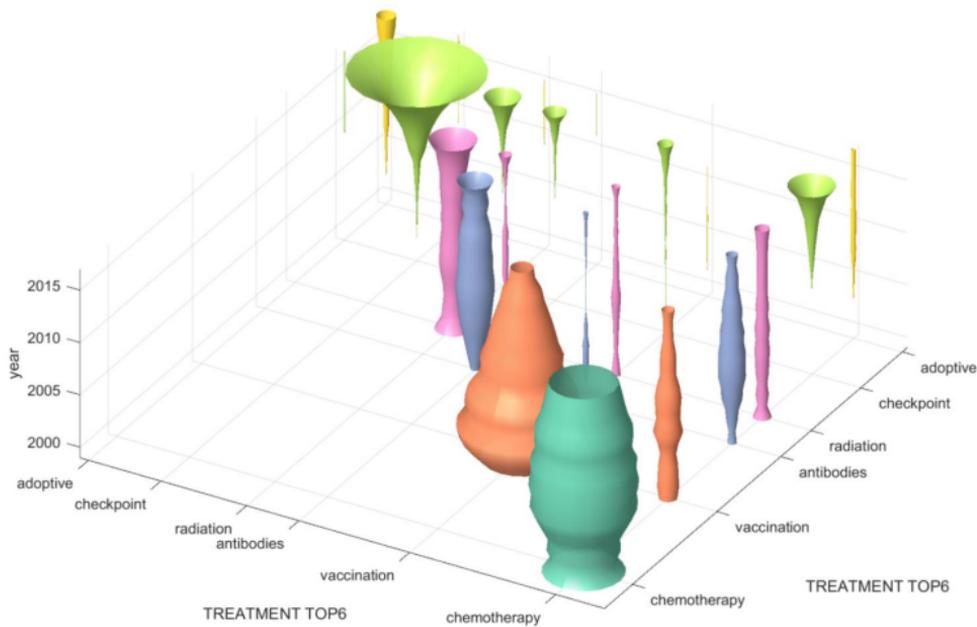
360 **Figure 6: Graph-based analysis of immune checkpoints.** In these graphs, the distance between  
361 two nodes denotes the co-occurrence while the color of the bubble denotes the frequency of  
362 occurrence (bubble sizes are log occurrence). (A) Before 2011, a cluster around  
363 CD80/CD86/CD40/CD28 dominated immune checkpoint research. (B) This has fundamentally  
364 changed since 2011: The field is now dominated by PD-1/PDL1, with CTLA-4 as a bystander. The  
365 number of relevant immune checkpoints has markedly increased. CD80 still occupies a central  
366 position in the network, linking the CD40/CD86/CD40L cluster with PD1/PD-L1/CTLA-4. TIM3  
367 and OX40 have also moved closer to the network's core, indicating an increasing importance  
368 despite few absolute hits.

369

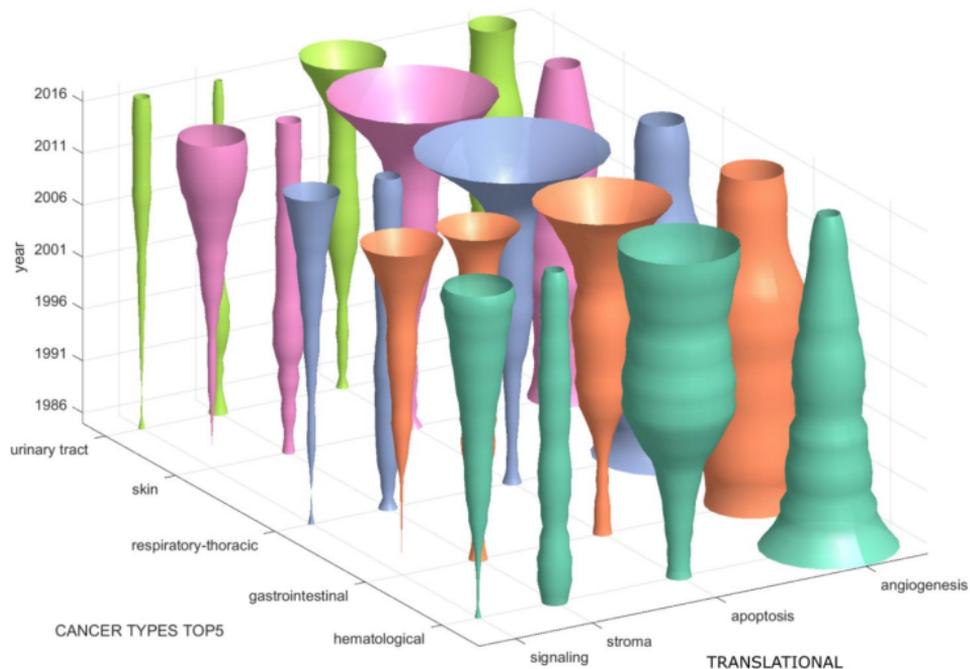


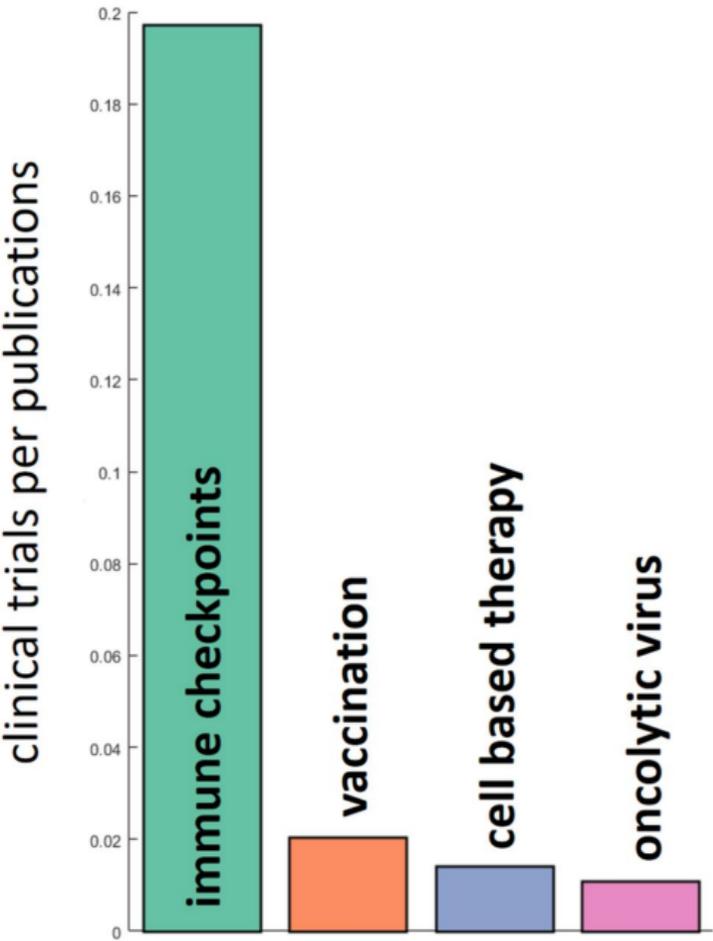
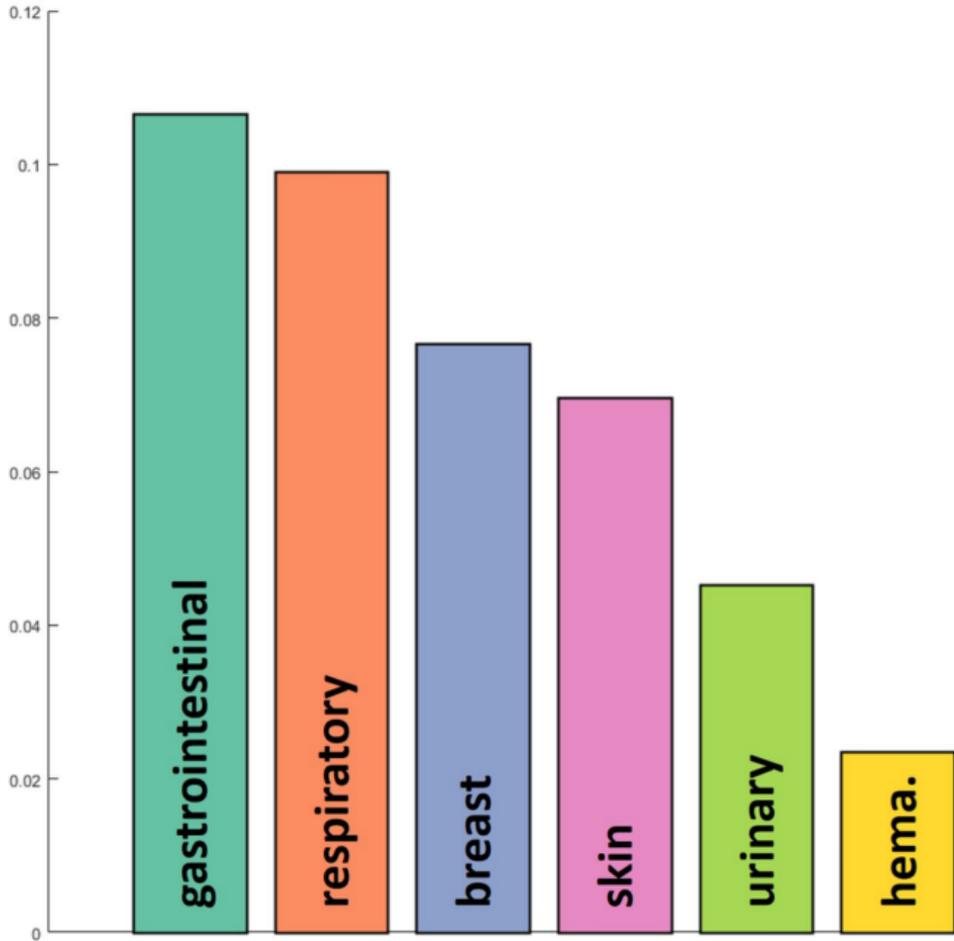


A

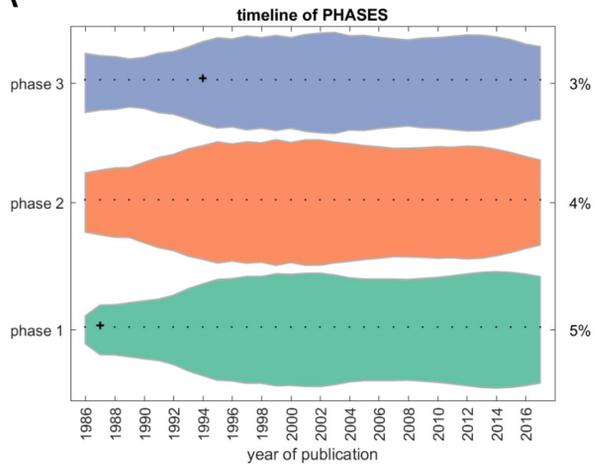


B

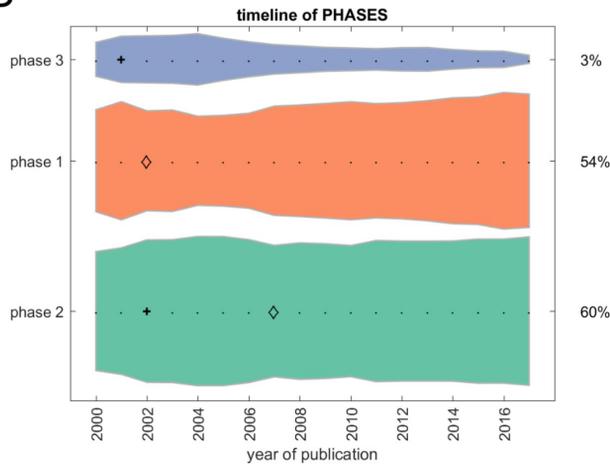


**A** translational efficiency / therapy types**B** translational efficiency / tumor types

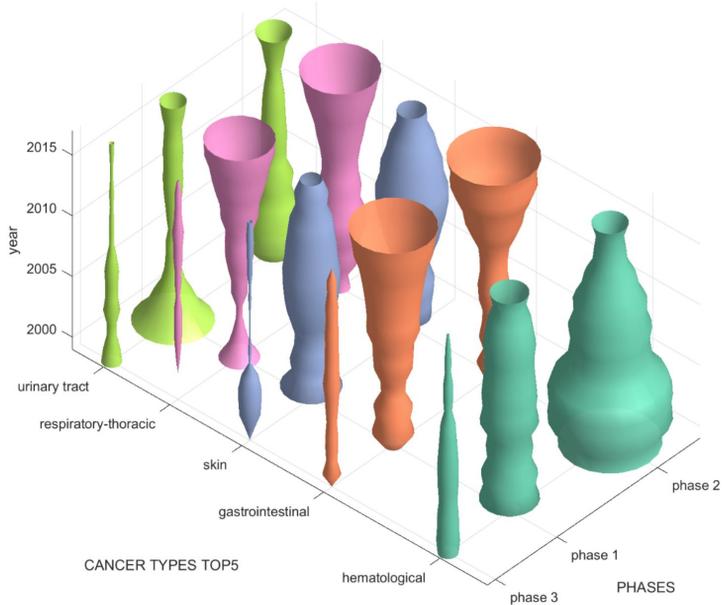
A

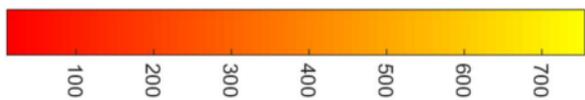
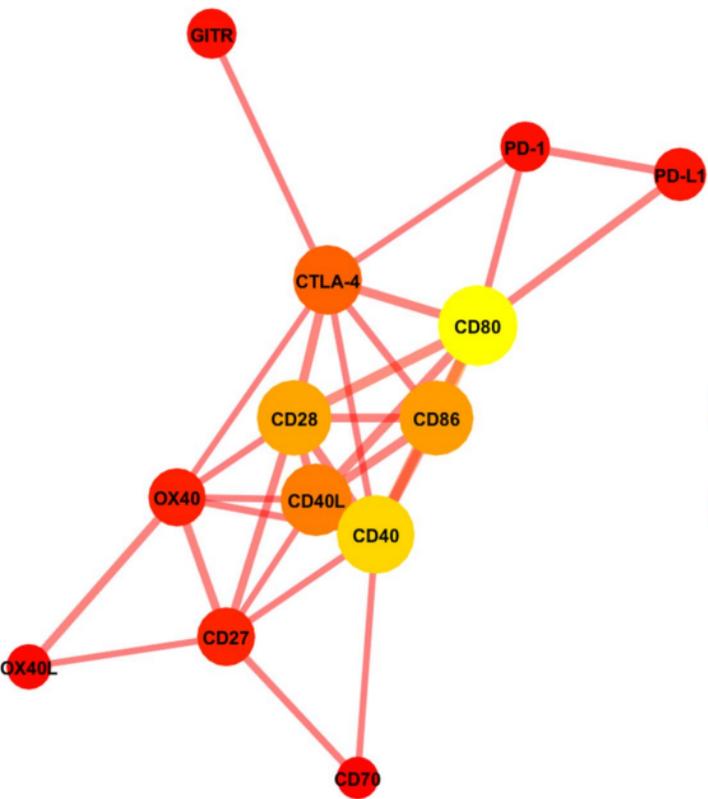


B



C



**A** checkpoints 1986-2010**B** checkpoints 2011-2016