Scientometrics (2020) 124:153–168
https://doi.org/10.1007/s11192-020-03429-1

Public–private contribution to biopharmaceutical discoveries: a bibliometric analysis of biomedical research in UK

Gianluca Fabiano[1] · **Andrea Marcellusi**[1] · **Giampiero Favato**[1]

Received: 16 July 2019 / Published online: 28 March 2020

Abstract
Basic research creates new knowledge that fuels technological advances. However, budgetary concerns and escalating R&D prices are challenging organizations to show returns from investments in scientific research. Few attempts are made to analyse research that leads to pharmaceutical innovation. In particular, the financial contribution of public and private organizations to the riskiest stage of biomedical discovery has remained unclear and partially unexplored. This study is a first attempt to shed light on the financial support to basic research by public and private sectors using publications data. We conducted an exploratory analysis of funding acknowledgments on publications authored by the founding scientists of 91 'drug originator' companies in United Kingdom. The nature and distribution of the support acknowledged to the research conducted before the company creation was analysed and the impact of publications and type of support were statistically tested. We found the majority of publications acknowledged public institutions, whereas, commercial organisations were likely to support those with privately affiliated authors. Based on these findings, we discussed the need to foster collaborative research and to set adequate incentives for shared risks and benefits from investments in knowledge creation.

Keywords Basic research · Drug discovery · Innovation · Funding acknowledgment

Introduction

The roles of public and private sectors in the discovery and development of new drugs are part of a long-running debate. In the past decades, the rise of novel collaborations in the form of industry-academy partnerships have blurred the traditional border between the two sectors (Etzkowitz 2003; Leydesdorff and Etzkowitz 1996; Bozeman et al. 2013; Bozeman 2000). Budgetary concerns and highly priced medicines are yielding increasing pressure on governments and raising new questions about the rates of return from public spending. Moreover, the argument that public sector is responsible for a large share of drug

✉ Gianluca Fabiano
Gia.fabiano@gmail.com

[1] Institute for Leadership and Management in Health, Kingston University London, Kingston Hill, Kingston upon Thames, London KT2 7LB, UK

development, have fuelled proposals to recoup profits from government funded drugs and to avoid tax payers to pay twice: first with taxes for publicly funded research, and then through monopoly of prices or restricted access (Sampat and Lichtenberg 2011).

The extent of public sector involvement in basic scientific research is recognized under the major role played by governmental funding agencies which affect the cognitive development of science and ultimately the economic growth and competitiveness (Kearnes and Wienroth 2009). However, the relation between funding input and research outputs is difficult to map due to the number of funding sources and the variety of collaborative relationships.

To probe collaboration patterns, the use of publication data has been prominently focused on authorship, as an indicator of scientific productivity, and citations as a measure of impact (Abramo et al. 2009). More recently, however, the study of the acknowledgments section of published literatures has taken part to so-called "reward triangle" and it is gaining ground as a source of data (Costas and van Leeuwen 2012; Cronin and Weaver 1995). In particular, acknowledgments of financial support are of prominent importance as they indicate the source of funding or economic support which made possible the research and the publication. Moreover, acknowledgments have the potential to map the funding inputs without the need to gain direct access to data via funders or researchers individually (Rigby 2011; Grassano et al. 2017). Accordingly, the study of acknowledgment can also be of interest to highlight research policy purposes (Álvarez-Bornstein et al. 2017; Paul-Hus et al. 2016).

Given the aforementioned potentials for using publications data, in this study we seek to improve our understanding on the financial aspects underlying the basic research which led to the discovery of pharmaceuticals. In particular, we explore the scientific production of scientists involved in the creation of biotechnology companies in United Kingdom (UK) and which originated a drug development project. Our objective is to identify the nature of the support acknowledged by the publications authored by the founders of such enterprises. We seek to shed light on the financial contribution of the public and private organizations to the science which anticipated the company creation. Doing so, we contribute to highlight some the financial levers which characterize the earliest and riskiest stage of pharmaceutical innovation.

The remainder of this paper is organised as follows. First, we present the conceptual motivation behind this study which we use to formulate a set of research questions. Next, we describe the study design and the methodology we applied to conduct our analysis. We subsequently present the results of our study and we discuss the main policy implications.

Previous literature and objectives

Since the biotech revolution, much has changed in the paradigm by which biomedical research and development (R&D) are conducted and theorized. The linear model of innovation is one of the first conceptual frameworks that was developed to understand the relation between science, technology and their application to the economy (Godin 2006). This approach defines innovation as causal process that leads from basic to applied research and development ending with production and diffusion (Godin 2006). Consistently with the literature that followed (Bush 2001) seminal work, much basic biomedical research is conducted in academic/non-profit institutions and is funded by the public sector (Zycher et al. 2010). Conversely, the bulk of applied and development science that concerns the

translation of the research findings into products or processes is funded mostly by the private sector. Public basic research is characterized by a high degree of uncertainty in both its scientific maturity and its potential market applicability (Toole 2007). Therefore, the private sector has inadequate incentives to invest in basic research (Galkina Cleary et al. 2018; Arrow 1978).

In the last decades, many criticisms concerning the linearity of the model have been raised and authors have claimed the inclusion of the dynamic interactions between academic and commercial sectors (Cockburn and Henderson 1996, 1998). Since the enactment of the Bayh–Dole Act in the USA in 1980, there has been a substantial rise in the commercialization of science and other forms of university technology transfer (Siegel and Wright 2015). Universities have broadened the traditional mission of educating students to include the patenting and commercialization of research discoveries (Bok 2009). As a result, multiple channels and processes of knowledge interaction have been established between universities and industry which have been featured in the growing literature on university–industry relations (Bekkers and Freitas 2008; Bercovitz and Feldman 2006; Bercovitz et al. 2001).

In this context, an increasing number of university scientists started for-profit companies to develop their scientific breakthroughs (Etzkowitz 1983; Shane and Khurana 2003; Stuart and Ding 2006). In particular, academic scientists have been of critical importance for the emergence of the biotechnology industry. Mansfield (1991) pointed out that the pharmaceutical industry had the highest percentage of new products based on academic research. Many start-ups have been either founded by scientists or heavily rely on work of public science (Zucker et al. 1998; McMillan et al. 2000; Audretsch and Stephan 1996). Stuart and Ding (2006) estimate that half of all biotechnology firms have been founded by university scientists, most of whom maintained academic appointments post-founding. Azoulay et al. (2007) found that 38% of the members of a random sample of 3800 U.S.-based, academic life scientists had, at some point in their careers, co-authored one or more papers with scientists working in the private-sector.

Past research has employed publications data to study company innovative performances. For example, using bibliographic information of researchers affiliated with 190 UK-based biotechnology firms, Jong and Slavova (2014) demonstrated that firms adopting open science strategies were more prominent in the R&D landscape with benefits related to the development of radical products innovations. Some empirical studies on scientific human capital have demonstrated the varying performance impact of a mix of scientists in terms of their orientation towards publishing and patenting (Baba et al. 2009; Stokes 2011; Subramanian et al. 2013; Zucker et al. 2002). For these reasons, published studies by founding scientists contain a rich source of information which can be used to probe probing collaboration patterns, interdisciplinary linkages and other research spill overs (Wang and Shapira 2015). This is confirmed by Sternitzke (2010) who reported that each drug is accompanied by, on average, about 19 journal publications and 23 additional patents.

In the past literature, some attempts to make use of publication data were made either through case studies or by studies which used grant proposals, CV or performance report as sources for funding information (Stuart and Ding 2006; Heinze et al. 2007). The acknowledgment field of publications has been used by a growing number of bibliometric studies. Until recent times, it was very difficult to carry out analyses on acknowledgments since this information was not available in bibliographic databases. However, since 2008, Web of Science (WoS) started including funding acknowledgment data and this opened up new possibilities for investigations. Some studies analysed the acknowledgment sections by disciplines or countries (Díaz-Faes and Bordons 2014; Wang and Shapira 2015;

Costas and van Leeuwen 2012). These highlighted the acknowledgment appeared more frequent in Anglo-American journals and in the so-called 'hard sciences' such as chemistry and psychology (Cronin et al. 2004). Other contributions measured the role of funded articles as an indicator of research quality (Gillett 1991; Costas and van Leeuwen 2012; Zhao 2010). Also, partnerships between the public and private sectors in multiple disciplinary areas have been studied using funding acknowledgement section of Spanish articles (Morillo 2016) founding that 79% was financed by the public and 18% in combination with private. The accuracy and completeness of the extraction process adopted by the extant literature have been the object of some recent contribution which also questioned and tested the robustness of the bibliographical databases reporting funding information (Rigby 2011; Grassano et al. 2017).

Yet, despite the potentials and the magnitude of publications as a source of data, the existing body of literature have rarely employed these into the analysis of the funding of pharmaceutical innovation. Based on the previous research, this paper is a first attempt to analyse the nature and the scope of the financial contribution acknowledged by the literature which lead to biopharmaceutical discoveries. To do so, we chose to analyse the funding acknowledgements (FA) in scientific publications authored by the founders of biopharmaceutical companies before the incorporation. Our focus is, therefore, the basic and applied research that is undertaken primarily for the advancement of scientific knowledge, with or without a specific practical application in view (Godin 2006). The main objectives addressed in this article are the following:

- To explore the nature of the financial support acknowledged in the publications by scientists involved in the foundation of biotechnology companies.
- To highlight the distribution of funding organizations across countries and disciplines.
- To analyse the extent to which funding information have higher impact in terms of citations.

Moreover, based on the results of the proposed objectives, we aim to critically assess the position of public and private sector towards the funding and of the earliest stage of pharmaceutical innovation and discuss potential research policy implications.

Methods

Sample

To construct our sample, we identified 91 UK-based biotechnology firms founded between 1996 and 2016, which originated at least one drug development project[1] and received venture capital (VC) support according to the GlobalData Pharma Intelligence Centre's database. We assumed that investments from VC represented a promising stage at which discoveries made in basic research may conduct and after which there might be further investments in development research.

[1] *Drug originator* is when the drug is "conceptualized, discovered and initially developed and the intellectual rights originate"; *Active* is when the developmental status or the product has been updated in the last 2 years, if in clinical development, or 4 years if the drug was in a preclinical development stage (GlobalData PlC).

Further, we collected the biographical information of the founding scientists (*full name, role, previous affiliation*). To do so, we accessed Beaurhurst database that contains key information on high-growth companies that have secured equity fundraising in UK. When founders' or primary scientists' personal details were not reported, we conducted a web search of companies' websites. For each company, multiple identities were selected and ranked according to their role. Specifically we gave more relevance to the scientists or those individuals that, accordingly to the information provided, were involved in the identification of the company science and its foundation (see "Appendix" for ranking criteria). The collection was conducted by two independent researchers, results have been compared afterword; any disagreement and risk of bias were solved through a discussion between the authors.

Data collection

A literature search was conducted using two search queries on different bibliographic databases.

First (method 1), we aimed to capture the most relevant publications by the founding scientists before the company creation. Accordingly, using Thomson Reuter's Web of Science (WoS) database, we collected the ten most cited articles published by the company founder before the year of the incorporation.

Method 1, search query author: "name of founding scientist"_ FILTER < (year of company foundation)

A second research query (method 2) was made associating the founders' identity with the information regarding the products in the company pipeline. Details of the companies' pipelines were downloaded from GlobalData database. Specifically, we looked at compounds at the most advanced stage in development in each company pipeline. Next, we collected information on: (a) the disease or condition for which the drug is or is seeking approval (*therapy area*); (b) the structural, biological, chemical and/or pharmaceutical features of the drug (*molecule type*); (c) the specific effect on the indication or symptoms that are intended to treat (*mechanism of action*).

Accordingly we downloaded the first publication appearing on Google Scholar for each of the included founding scientists. Scholar is recognized as one of the few academic search engines that combines several approaches in a single algorithm as a result of weighing the full text, the author, the publication in which the articles appears and how often the piece has been cited in other scholarly literature (Beel and Gipp 2009). This approach allowed the identification of the most relevant article per author according to the company science.

Method 2, search query "Mechanism of action "AND/OR" molecule type "AND/OR "therapy area" AND author: "name of scientific founder(s)"_ FILTER < (year of company foundation)

As a result of the two collecting strategies we aimed to include the most influential publications in the founders' backgrounds (method 1) including the research made in the area that was then carried on within the enterprise as the base for the drug development (method 2).

Overall, we collected only articles written in English and published in peer-reviewed scientific journals. Meeting abstracts, commentaries and reviews were excluded. This two-steps procedure yielded 714 results (620 with method 1 and 94 using method 2). Duplicates originating from the two research methods (77 publications, 11%) were removed.

Data analysis

Data cleansing was required to extract and harmonize the information contained in our set of publications. A first problem concerned the standardisation and disambiguation of author names. This arises when several authors share the same name, but also one author expresses his/her name in different ways. As pointed out by D'Angelo et al. (2011), manual inspection to disambiguate authors can be very effective for small population of scientists and data integration from external sources can be very helpful to ease the task. In our work, we used the information collected about the founders' background from companies' websites to make an efficient use of the "article group" option provided by WoS.

Secondly, our sample was checked to identify funding information contained in the acknowledgement section of each article. Due to the reduced size of our sample, we opted for extracting, coding and interpreting funding data in addition to WoS manually. Accordingly to Grassano et al. (2017), when there was no clear way to detect the financial nature of the support, we adopted a conservative approach and the acknowledged organisation was not included in the list of funders. Also, as the intention of our work was to highlight the financial contribution supporting the research, we included only funding data that were explicitly reported to the founder or acknowledged to the study, with the exclusion of all funding provided explicitly to other authors. Once the relevant funding information for a given publication were identified, we extracted the full name of the organisation and added the code of the country where this was located and labelled according to the scope (public, private). Governmental bodies, charities and non-profit organisations were grouped as 'public', whereas private foundations held by commercial organisations were included under 'private' (Sussex et al. 2016; Viergever and Hendriks 2016). A thesaurus of the various names and acronyms of funding agencies was created.

Thirdly, affiliation analysis was conducted following the approach applied by Hottenrott and Lawson (2017). Specifically, each author on the selected publications was attributed one or more institution based on the [C1] address field in WoS. All addresses were coded by institution type and assigned to a unique code. This was conducted semi-manually, meaning that a search algorithm containing word elements such as "univ", "hosp", "ltd" was undertaken using Excel, and these were marketed as universities, hospitals or companies respectively. The retrieved institutions were finally coded as belonging to the higher education sector (HEI),[2] public or semi-governmental research institutes (PRO), private companies (PRIV).

The bibliometric analysis of the acknowledgement section as well as authors affiliations was performed in Excel. The Statistical Package for the Social Sciences (SPSS, version 24.0 for Windows) and STATA were used to perform statistical analysis and test relations in our dataset. Figure 1 illustrates the study design, inclusion criteria and data collection strategy.

[2] The majority of hospitals were university teaching hospitals and thus assigned to HEI (Viergever and Hendriks 2016).

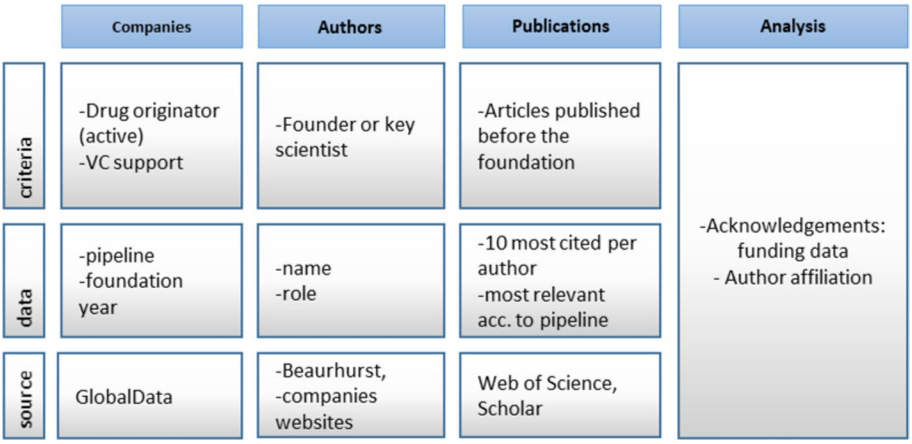

Fig. 1 Study design

	n *(%)*
Included companies	91
Founder scientists, authors	141
Publications (Method 1)	620
Publications (Method 2)	94
Tot publications	714
Excluded (eg: duplicates)	77
Tot included publications	**637** ***(89%)***
Publications reporting FA	**403** ***(63%)***
Citations per publication*	239.7
Authors per publication*	7.7
Articles per company*	5
Funding organisations per publication*	1.8

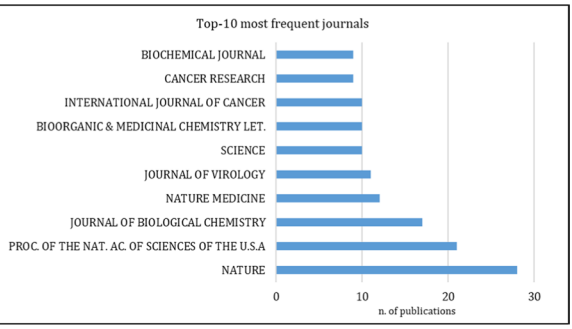

Fig. 2 Descriptive results

Results

A total of 637 publications, made between 1970 and 2015 and authored by 141 founders of 91 biotechnology companies was included in our study. Funding data addressed to the founder/scientist or generically acknowledged to the study was disclosed in 403 (63%) articles (Fig. 2).

The support from public institutions was reported in 329 (82%) publications whereas that from the industry or a combination of private and public organisations ("mix") was found in 28 (7%) and 46 (11%) respectively. Authors were affiliated to the higher education sector (HEI) in 72% of our sample, 14% to public or semi-governmental research institutes (PRO) and 10% to private companies (PRIV); 3% showed multiple affiliations (Fig. 3).

Table 1 illustrates the number of publications by type of funder and authors' affiliation. When the publications were privately financed, authors were likely to be affiliated with private organisations (52% vs 6% of publicly financed). A Fisher's exact test of independence was performed and a significant relation between type of funding and author affiliation was found ($p < .001$). Overall, these results suggest that the distribution of funding data between

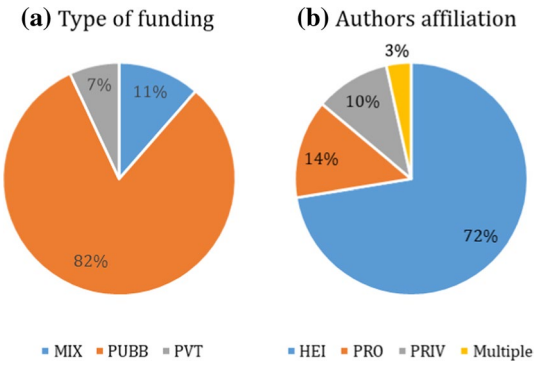

Fig. 3 Number of publications by type of funding **a** and authors affiliation **b**

Table 1 Funding to publications, affiliations

Affiliation	Public funding (%)	Private funding (%)	Mix (%)	Total (%)
HEI	251 (76%)	14 (48%)	28 (61%)	292 (72%)
PRO	47 (14%)	–	8 (17%)	55 (14%)
PRIV	17 (6%)	15 (52%)	10 (22%)	42 (10%)
Multiple	14 (4%)	–	–	14 (4%)
Total	329	29	46	403

HEI higher education institutions, *PRO* public or semi-governmental research institutes, *PRIV* private companies; multiple affiliations: HEI, PRO, PRIV or HEI, PRO, or HEI, PRIV

affiliations is not casual and authors belonging to universities or to private corporations are more likely to attract public or private funding respectively ($p < .001$).

The study also demonstrates as a myriad of different organizations are involved in the funding of biomedical research that leads to the creation of biotechnology enterprises in UK. As such, 261 different funders (206 public, 55 private) were individuated. Major funders, defined as those appearing in at least 10 publications, appeared in 70% of our sample showing that the majority of biomedical research was funded by a small number of organisations. Among public funders, the Wellcome Trust, Medical Research Council and the National Institute of Health (NIH) were the most acknowledged together with Cancer Research UK. The studies in our sample were supported by 77 public institutions based in UK that compared in 256 publications (63%). Also the contribution of 59 US funding organisations was acknowledged (24%) together with that of organisations from 22 different nationalities.

Funding acknowledgements also enabled the exploration of the extent to which funders are jointly disclosed in publications. In Fig. 4c we examine the co-occurrence of the ten major public funders. The analysis revealed that public funders most frequently jointly acknowledged were the Medical Research Council together with the Wellcome Trust or Cancer Research UK. Moreover, it is worth noting that not in a small proportion (22%), funding organisations were the same as the author affiliations and this value could be higher if we included all other authors' affiliation in the check in addition to founders.

The WoS category classification enabled to compare the content of the studies with the reported funding organisation and to analyse the topics that public and private funders mostly support. The areas with most publications were biochemistry, oncology and science and technology. Figure 5a, b profile public and private funders in terms of research domain

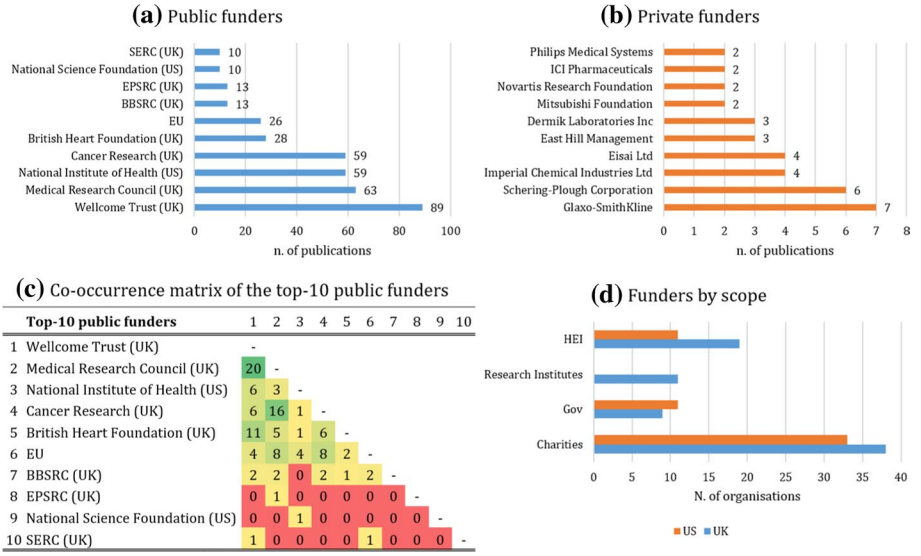

Fig. 4 Top-10 public **a** and private funders **b**, co-occurrence matrix **c** and scope of funders **d**

relative to the number of publications containing funding data. All major funders support was primarily focused on biochemistry and molecular biology. The Wellcome Trust together with the Medical Research Council showed a focus on science and technology and immunology. Cell biology was mostly investigated by Cancer Research UK that also showed a pronounced attention towards scientific production in oncology.

Further bibliometric indicators were used to gain a deeper understanding of how the interaction between public and private sectors take place within the production of scientific knowledge that leads primary scientists to the creation of biotechnology companies. Assuming the informal sense that the first and last author positions are generally perceived as the 'key' positions on a paper, the analysis demonstrated that when authors were publicly affiliated they were more often reported as first or last authors (80%) than private (72%). Overall, founders/scientists appeared first or last in byline position in 57% of our sample demonstrating a substantial contribution to the published work.

The impact of publications as measured by the number of citations per year was also assessed by applying non-parametric tests. Specifically, a Mann–Whitney U test showed that the number of citations per year was greater for publications with funding acknowledgements than for the articles that did not disclosed any funding to the author or the study ($p < .001$). On the other hand, a Kruskal–Wallis H test ($p = .148$) indicated that the distribution of citations per year was not statistically different across publications that received public funding compared to private or a mix of both sectors.

Discussion

The contribution of public and private entities to the early stages of pharmaceutical innovation was the object of this paper. We highlighted the public sector plays a key role in supporting the riskiest stage of pharmaceutical innovation that is when basic, blue-sky

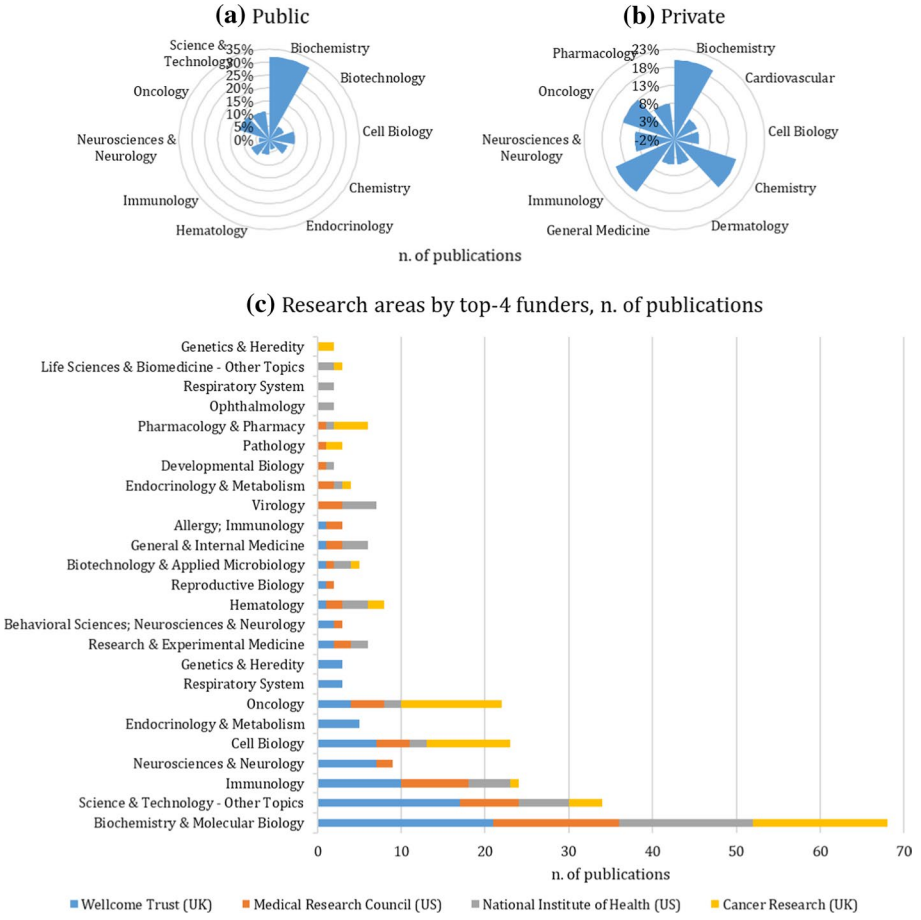

Fig. 5 Research areas supported by public **a**, private or mix **b**, top-4 major funders **c**

research is conducted with no practical or economical application in view. The work found that when financial acknowledgments are disclosed, the number of citations is higher and therefore the broader the impact of the publication. As a consequence, we may derive the basic research, if conducted with the support by public or private organisation has better chances to be the ground for future innovations.

The topic has been investigated through a bottom-up approach taking in analysis the publications made by leading scientists figuring as the founders of 91 biotechnology companies headquartered in UK. The discovery of innovative pharmaceuticals was explored assuming the production of scientific knowledge made before the incorporation to play a key role in the identification of the potentials for business. Discoveries made in basic research are significant for the process that leads to innovation. Specifically, the technology or knowledge transfer defined as the "movement of know-how, technical knowledge or technology from one organizational setting to another"(Bozeman 2000), plays an increasingly important role in stimulating economic development (Siegel et al. 2007). In this context, it has been recognized the research support from government or industry could play as catalyst for enabling the innovation process. However, despite the growing emphasis

on formalizing the interaction that take place between the higher educational system and industry as a vehicle to enhance innovation, the financial aspects related to this phenomenon were partially unexplored. As such, we analysed the type of support provided to the studies upon which we assumed the authors, once identified the potential of the results achieved, have started a commercial business. Contrarily we did not aim to assess or map the support given to biomedical research overall. Doing so, the current study sheds light on some financial aspects related to the discovery of pharmaceuticals.

Specifically, here we demonstrate the public sector plays a fundamental role in the financing of biopharmaceutical discoveries and so the risks associated. Using bibliographic information as a proxy, we evidenced the major role of the public sector as reported in the acknowledgments sections and through the affiliation of most of the authors. Surprisingly, the study found a tight relation between the type of support provided to publications and authors' affiliation. In particular, support from private sector was almost 4 times higher when the founders were already privately affiliated and lower (− 38%) when these, before starting their own company, were belonging to public or non-profit organisations. This was also evidenced for those studies financed by public institutions in which, however, the relation was weaker. Also, we demonstrated that the impact of publications that acknowledged financial support either by the public, private or a mix of both sector was higher in terms of annual citations. Contrarily, the type of funding did not seem to influence the number of yearly citations. Furthermore, we have examined the extent to which different organisations are acknowledged as funders of scientific publications. As a result of our investigation, we found a scientific landscape populated by a large number of different institutions. The distribution among publications was highly skewed meaning that few organisations financed the majority (70%) of the sample. Moreover, it is worth to notice that the support given by public institutions was more focused on broader areas of investigation such as biochemistry, cell biology whereas private organisations were contributing to more applied topics such as oncology, immunology and chemistry.

Furthermore, the application of bibliometric analysis to publication data have shown potentials to inform policy making on research funding. Here, the retrieved publication were used to profile the research portfolios of major public funders. As a result, the number of studies in oncology published with the support of Cancer Research UK (CRUK) were three times those of other institutions and constituted 20% of CRUK portfolio. Also, a specific focus on neurosciences and neurology was found in publications that disclosed the Wellcome Trust as supporting institution. In addition, as a result of the manual coding of the retrieved literature, the study measured the precision of WoS in reporting the information contained in the acknowledgement section. Consistently with the previous literature on this topic, we found that WoS did not report funding data for 20 publications (3%) for which our manual check evidenced financial support (false negatives).

In this study, we provide evidence whilst fully aware of the complexity of attributing a single conceptual framework to this topic and the limitations related to the use of bibliometric techniques in this field (Reichert and Milne 2002; Cockburn and Henderson 1996). A first critical aspect is associated with the manual coding of acknowledgements paratexts required to gain the appropriate understanding of the funding information to authors. Certainty the inclusion of the funding text [FU] in WoS since 2008 has enabled the analysis of acknowledgement section as the "*third edge*" of the reward triangle in bibliometric studies together with authorships and citations (Costas and van Leeuwen 2012).

However, the incompleteness of data provided by WoS and the lack of precision as highlighted by Grassano et al. (2017), have partially limited the use of this source. In addition, as said, our investigation was focused solely on authors that we identified as

the founder of the biotechnology companies. Hence, all funding information attributed to other authors were due to be excluded and from this we opted for a manual coding of the acknowledgement sections. Furthermore there were cases in which acknowledged information were not reported as "financial support" or "funding" but simply "supported by" and these were interpreted as financial as well. Consequently to that, the number of financed studies included in the study may be overestimated. Another limitation is represented by the way the identities of authors and founders have been identified and matched. The inclusion, despite it was made through a double-blind process conducted by two independent researchers, involved judgments and was subject to the reliability of available data. Precisely, we relied on the information contained in the accessed databases and companies' websites to determine whether the name of any of the present or past employer was reported as the founder or the leading scientist. Other names, of other people possibly involved in the creation of such enterprises may have been omitted and thus not included in the analysis. Also, when retrieving publications from the online databases, a number of homonymies were found and therefore limitations in associating names with titles were due to the required process of disambiguation.

In conclusion, although the study covers an extensive period of time, the recent boost in public–private efforts, such as the Innovative Medicine Initiative (IMI) at European level, may not be fully evidenced in our data. The IMI is a public–private partnership between the European Union and the European Federation of Pharmaceutical Industries which have produced more than four thousand publications involving collaboration between researchers across different sectors, institutions and countries (Laverty and Meulien 2019). Our study does however extend the extant literature on academic founders by focusing on scientific founders of biotechnology companies. By including different career backgrounds, our study demonstrates that company founders act as boundary-spanning 'gatekeepers' who facilitate access to socially embedded knowledge (Gittelman and Kogut 2003).

Future research may seek to clarify drivers of collaborative research by taking into account multiple processes of knowledge transfer (Fabiano et al. 2020). Research efforts are also encouraged to keep track of the funding given to applied research, for example by including publications made after the company foundation.

Conclusions

In this study, we traced the relationship between knowledge and innovation by taking a bottom up approach to identify the origin of an idea. Rather than observing research budgets top-down, we focused our analysis on scientists who published and started a company. By looking into this highly specific field of knowledge production, our results confirm the critical role played by non-profit research institutions in the financing of innovation.

Given the scale and the complexity of the challenges health systems are facing, the results of our study support the thesis that further polarization may occur if collaborative initiatives at research level are not encouraged. Collaboration at pre-competitive stage is a key aspect required to accelerate the flow of knowledge between university and industry. Therefore, we welcome policy initiatives and research efforts bringing multidisciplinary and multi-stakeholders together.

Furthermore, given the key role governments play a in both promoting innovation and in the generation of new market opportunities, risks and benefits need to be shared amongst actors at industrial level (Mazzucato 2011). From a policy perspective, the outcome of our

research advocates a shift towards a mission-oriented approach in which governments set the direction of innovation towards public needs. To support this, the pricing structure of pharmaceuticals should reflect the dual role taxpayers play in funding early innovation and as final users.

Funding This research did not receive any specific Grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix

See Table 2.

Table 2 Ranking criteria adopted to identify key employers

Ranking	Role
1	Co-Founder/inventor of the innovation
1	Scientific Founder
1	Scientific co-founder
1	Founder/Chief Scientist
1	Scientific Founder and Advisor
1	Founding Scientist and Head of department
2	Founder
2	Founder, CEO
2	Founder, CSO
2	Founder, Executive Director
2	Inventor/manager
3	Co-Founder
3	Co-Founder and Chief Medical Officer
3	Co-Founder and Chief Scientific Officer
3	Co-Founder, Chairman
3	Co-Founder, Chief Scientific Officer
3	Co-Founder, Managing Director
4	Non-Executive Director, Co-founder
4	Development
4	Director of medical, chemistry department
4	Chief Technology Officer
4	Chairman and Director of Drug

References

Abramo, G., D'Angelo, C. A., & Caprasecca, A. (2009). Allocative efficiency in public research funding: Can bibliometrics help? *Research Policy, 38*(1), 206–215.

Álvarez-Bornstein, B., Morillo, F., & Bordons, M. (2017). Funding acknowledgments in the web of science: Completeness and accuracy of collected data. *Scientometrics, 112*(3), 1793–1812.

Arrow, K. J. (1978). Uncertainty and the welfare economics of medical care. In J. D. Hey (Ed.), *Uncertainty in economics* (pp. 345–375). Amsterdam: Elsevier.

Audretsch, D. B., & Stephan, P. E. (1996). Company-scientist locational links: The case of biotechnology. *The American Economic Review, 86*(3), 641–652.

Azoulay, P., Ding, W., & Stuart, T. (2007). The determinants of faculty patenting behavior: Demographics or opportunities? *Journal of Economic Behavior & Organization, 63*(4), 599–623.

Baba, Y., Shichijo, N., & Sedita, S. R. (2009). How do collaborations with universities affect firms' innovative performance? The role of "Pasteur scientists" in the advanced materials field. *Research Policy, 38*(5), 756–764.

Beel, J., & Gipp, B. (2009). Google Scholar's ranking algorithm: an introductory overview. In *Proceedings of the 12th international conference on scientometrics and informetrics (ISSI'09)* (Vol. 1).

Bekkers, R., & Freitas, I. M. B. (2008). Analysing knowledge transfer channels between universities and industry: To what degree do sectors also matter? *Research Policy, 37*(10), 1837–1853.

Bercovitz, J., & Feldman, M. (2006). Entpreprenerial universities and technology transfer: A conceptual framework for understanding knowledge-based economic development. *The Journal of Technology Transfer, 31*(1), 175–188.

Bercovitz, J., Feldman, M., Feller, I., & Burton, R. (2001). Organizational structure as a determinant of academic patent and licensing behavior: An exploratory study of Duke, Johns Hopkins, and Pennsylvania State Universities. *The Journal of Technology Transfer, 26*(1–2), 21–35.

Bok, D. (2009). *Universities in the marketplace: The commercialization of higher education*. Princeton: Princeton University Press.

Bozeman, B. (2000). Technology transfer and public policy: A review of research and theory. *Research Policy, 29*(4–5), 627–655.

Bozeman, B., Fay, D., & Slade, C. P. (2013). Research collaboration in universities and academic entrepreneurship: The-state-of-the-art. *The Journal of Technology Transfer, 38*(1), 1–67.

Bush, V. (2001) *Science: The endless frontier.* ACLS History E-Book Project.

Cockburn, I., & Henderson, R. (1996). Public-private interaction in pharmaceutical research. *Proceedings of the National Academy of Sciences of the United States of America, 93*(23), 12725–12730.

Cockburn, I. M., & Henderson, R. M. (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *The Journal of Industrial Economics, 46*(2), 157–182.

Costas, R., & van Leeuwen, T. N. (2012). Approaching the "reward triangle": General analysis of the presence of funding acknowledgments and "peer interactive communication" in scientific publications. *Journal of the American Society for Information Science and Technology, 63*(8), 1647–1661.

Cronin, B., Shaw, D., & Barre, K. L. (2004). Visible, less visible, and invisible work: Patterns of collaboration in 20th century chemistry. *Journal of the American Society for Information Science and Technology, 55*(2), 160–168.

Cronin, B., & Weaver, S. (1995). The praxis of acknowledgement: From bibliometrics to influmetrics. *Revista Española De Documentación Científica, 18*(2), 172.

D'Angelo, C. A., Giuffrida, C., & Abramo, G. (2011). A heuristic approach to author name disambiguation in bibliometrics databases for large-scale research assessments. *Journal of the American Society for Information Science and Technology, 62*(2), 257–269.

Díaz-Faes, A. A., & Bordons, M. (2014). Acknowledgments in scientific publications: Presence in Spanish science and text patterns across disciplines. *Journal of the Association for Information Science and Technology, 65*(9), 1834–1849.

Etzkowitz, H. (1983). Entrepreneurial scientists and entrepreneurial universities in American academic science. *Minerva, 21*(2), 198–233.

Etzkowitz, H. (2003). Innovation in innovation: The triple helix of university-industry-government relations. *Social Science Information, 42*(3), 293–337.

Fabiano, G., Marcellusi, A., & Favato, G. (2020). 'Channels and processes of knowledge transfer: How does knowledge move between university and industry? *Science and Public Policy*. https://doi.org/10.1093/scipol/scaa002.

Galkina Cleary, E., Beierlein, J. M., Khanuja, N. S., McNamee, L. M., & Ledley, F. D. (2018). Contribution of NIH funding to new drug approvals 2010–2016. *Proceedings of the National Academy of Sciences of the United States of America, 115*(10), 2329–2334.

Gillett, R. (1991). Pitfalls in assessing research performance by grant income. *Scientometrics, 22*(2), 253–263.
Gittelman, M., & Kogut, B. (2003). Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science, 49*(4), 366–382.
Godin, B. (2006). The linear model of innovation: The historical construction of an analytical framework. *Science, Technology and Human Values, 31*(6), 639–667.
Grassano, N., Rotolo, D., Hutton, J., Lang, F., & Hopkins, M. M. (2017). Funding data from publication acknowledgments: Coverage, uses, and limitations. *Journal of the Association for Information Science and Technology, 68*(4), 999–1017.
Heinze, T., Shapira, P., Senker, J., & Kuhlmann, S. (2007). Identifying creative research accomplishments: Methodology and results for nanotechnology and human genetics. *Scientometrics, 70*(1), 125–152.
Hottenrott, H., & Lawson, C. (2017). A first look at multiple institutional affiliations: A study of authors in Germany, Japan and the UK. *Scientometrics, 111*(1), 285–295.
Jong, S., & Slavova, K. (2014). When publications lead to products: The open science conundrum in new product development. *Research Policy, 43*(4), 645–654.
Kearnes, M. & Wienroth, M. (2009). 'Arm's length'?: Narratives of impact and autonomy in UK Research Councils.
Laverty, H., & Meulien, P. (2019). The innovative medicines initiative—10 years of public-private collaboration. *Frontiers in Medicine, 6,* 275.
Leydesdorff, L., & Etzkowitz, H. (1996). Emergence of a Triple Helix of university—Industry—Government relations. *Science and Public Policy, 23*(5), 279–286.
Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy, 20*(1), 1–12.
Mazzucato, M. (2011). The entrepreneurial state. *Soundings, 49*(49), 131–142.
McMillan, G. S., Narin, F., & Deeds, D. L. (2000). An analysis of the critical role of public science in innovation: The case of biotechnology. *Research Policy, 29*(1), 1–8.
Morillo, F. (2016). Public–private interactions reflected through the funding acknowledgements. *Scientometrics, 108*(3), 1193–1204.
Paul-Hus, A., Desrochers, N., & Costas, R. (2016). Characterization, description, and considerations for the use of funding acknowledgement data in web of science. *Scientometrics, 108*(1), 167–182.
Reichert, J. M., & Milne, C. (2002). Public and private sector contributions to the discovery and development of "impact" drugs. *American Journal of Therapeutics, 9*(6), 543–555.
Rigby, J. (2011). Systematic grant and funding body acknowledgement data for publications: New dimensions and new controversies for research policy and evaluation. *Research Evaluation, 20*(5), 365–375.
Sampat, B. N., & Lichtenberg, F. R. (2011). What are the respective roles of the public and private sectors in pharmaceutical innovation? *Health Affairs, 30*(2), 332–339.
Shane, S., & Khurana, R. (2003). Bringing individuals back in: The effects of career experience on new firm founding. *Industrial and Corporate Change, 12*(3), 519–543.
Siegel, D. S., Veugelers, R., & Wright, M. (2007). Technology transfer offices and commercialization of university intellectual property: Performance and policy implications. *Oxford Review of Economic Policy, 23*(4), 640–660.
Siegel, D. S., & Wright, M. (2015). Academic entrepreneurship: Time for a rethink? *British Journal of Management, 26*(4), 582–595.
Sternitzke, C. (2010). Knowledge sources, patent protection, and commercialization of pharmaceutical innovations. *Research Policy, 39*(6), 810–821.
Stokes, D. E. (2011). *Pasteur's quadrant: Basic science and technological innovation*. Washington: Brookings Institution Press.
Stuart, T. E., & Ding, W. W. (2006). When do scientists become entrepreneurs? The social structural antecedents of commercial activity in the academic life sciences. *American Journal of Sociology, 112*(1), 97–144.
Subramanian, A. M., Lim, K., & Soh, P. (2013). When birds of a feather don't flock together: Different scientists and the roles they play in biotech R&D alliances. *Research Policy, 42*(3), 595–612.
Sussex, J., Feng, Y., Mestre-Ferrandiz, J., Pistollato, M., Hafner, M., Burridge, P., et al. (2016). Quantifying the economic impact of government and charity funding of medical research on private research and development funding in the United Kingdom. *BMC Medicine, 14*(1), 32.
Toole, A. A. (2007). Does public scientific research complement private investment in research and development in the pharmaceutical industry? *The Journal of Law and Economics, 50*(1), 81–104.
Viergever, R. F., & Hendriks, T. C. (2016). The 10 largest public and philanthropic funders of health research in the world: What they fund and how they distribute their funds. *Health Research Policy and Systems, 14*(1), 12.
Wang, J., & Shapira, P. (2015). Is there a relationship between research sponsorship and publication impact? An analysis of funding acknowledgments in nanotechnology papers. *PLoS ONE, 10*(2), e0117727.

Zhao, D. (2010). Characteristics and impact of grant-funded research: A case study of the library and information science field. *Scientometrics, 84*(2), 293–306.
Zucker, L. G., Darby, M. R., & Armstrong, J. (1998). Geographically localized knowledge: Spillovers or markets? *Economic Inquiry, 36*(1), 65–86.
Zucker, L. G., Darby, M. R., & Armstrong, J. S. (2002). Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. *Management Science, 48*(1), 138–153.
Zycher, B., DiMasi, J. A., & Milne, C. (2010). Private sector contributions to pharmaceutical science: Thirty-five summary case histories. *American Journal of Therapeutics, 17*(1), 101–120.