Research Policy 53 (2024) 104927

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Geography of authorship: How geography shapes authorship attribution in big team science

Jarno Hoekman [a,*], Bastian Rake [b]

[a] Innovation Studies, Copernicus Institute of Sustainable Development, Utrecht University, the Netherlands
[b] School of Business, Maynooth University, Maynooth, Ireland

ARTICLE INFO

Keywords:
Authorship
Geography of science
Team science
Clinical trials
Scientific collaboration

ABSTRACT

The steady growth of large geographically dispersed research projects challenges existing norms for authorship attribution and has raised concerns over global inequalities in authorship opportunities. This paper therefore examines how geography plays a role in authorship attribution to local researchers that contribute to large scientific teams from various cities across the globe. We develop theory that considers how authorship opportunities for local researchers may vary depending on how they are spatially embedded in projects and the local resources they draw upon. We empirically apply this framework to the context of multi-city clinical trials where a common authorship challenge concerns the attribution of site investigators on publications. To account for selection effects in our empirical set-up, we estimate authorship likelihood conditional on data collection contributions. Our results show that authorship likelihoods differ considerably across research projects and cities. We observe that, after controlling for project characteristics, authorship likelihoods are higher when local site investigators are located in cities that are geographically proximate to coordinating sponsors and when they face less national competition. We also find that local scientific reputation and the extent to which project contributions are directed to local problems are positively related to authorship likelihood. Observed findings are markedly more pronounced for industry-sponsored versus publicly-sponsored trials and when attributing authorship to a lead author compared to any author. Based on these findings, we discuss various ways through which authorship policies and initiatives could foster equitable authorship opportunities in large teams independent of location and as a fundamental principle for the conduct of science.

1. Introduction

The increasing complexity of scientific and societal problems, a rapidly advancing knowledge frontier and a decline in travel and communication costs have gradually turned science into a highly collaborative team activity (Hall et al., 2018; Katz and Martin, 1997; Wuchty et al., 2007). Over the last decades, research teams have been growing in almost all fields and it is no longer uncommon to find big teams with >100 contributors and research papers with 10 or more authors (Milojević, 2014; Wuchty et al., 2007). Big research teams particularly play an important role in the development of existing ideas (Wu et al., 2019), often by working closely with non-scientific actors in context-driven modes of knowledge production that transgress disciplinary boundaries (Beck et al., 2022; Gibbons et al., 1994; Hessels et al., 2009). Their growth is accompanied by an increasing bureaucratic organisation of work (Walsh and Lee, 2015) and a changing geography of science, as demonstrated by a sharp increase in international research collaboration and involvement of more and more cities in scientific knowledge production (Adams et al., 2005; Csomós, 2018; Hoekman et al., 2010; Jones et al., 2008).

The rise of big geographically dispersed team science is challenging existing institutions for rewarding researchers' contributions in a research project with authorship. While this is far from a new problem (cf. Zuckerman, 1968), the growth of large research teams is making it more difficult to assess the contributions that individual researchers make to a project, including when such contributions are substantive enough to merit authorship (Haeussler and Sauermann, 2013; Jabbehdari and Walsh, 2017). The issue is particularly salient for interdisciplinary research projects that are characterized by greater division of labour and fragmentation of tasks (Haeussler and Sauermann, 2020;

* Corresponding author at: Innovation Studies, Copernicus Institute of Sustainable Development, Faculty of Geosciences, Utrecht University, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands.
E-mail addresses: j.hoekman@uu.nl (J. Hoekman), bastian.rake@mu.ie (B. Rake).

https://doi.org/10.1016/j.respol.2023.104927
Received 23 December 2022; Received in revised form 28 September 2023; Accepted 20 November 2023
Available online 22 December 2023

Walsh and Lee, 2015). Institutionally, authorship guidelines such as those proposed by the International Committee for Medical Journal Editors provide guidance on who should qualify for authorship but they have been criticized for providing considerable openings for interpretation and not being fit-for-purpose for interdisciplinary research projects (Allen et al., 2014; McNutt et al., 2018; Sauermann and Stephan, 2013; Wager, 2009). In practice, authorship attribution in research projects therefore tend to be shaped by a combination of formal guideline interpretations and informal authorship norms and conventions established in specific scientific communities and disciplines (Haeussler and Sauermann, 2020; Jabbehdari and Walsh, 2017; Laudel, 2002; Sauermann and Haeussler, 2017; Wager, 2009).

Given the interpretative ambiguities of authorship guidelines, scholars have started to examine how social factors play a role in authorship attribution,[1] independent from project contributions. Studies have shown that authorship attribution in research teams is subject to negotiation processes and dependent on individual bargaining power (Lissoni et al., 2020, 2013) as well as on project and field-level structures in which researchers are embedded (Jabbehdari and Walsh, 2017; Marušić et al., 2011). On the individual level, prior research performance and scientific eminence are positively related with authorship attribution once controlled for similarity in contribution (Haeussler and Sauermann, 2013; Larivière et al., 2016; Lissoni et al., 2013, 2020). On the project level, there is evidence that authorship attribution is influenced by research team size and whether projects are conducted at a single or multiple sites (Jabbehdari and Walsh, 2017).

While prior studies have thus started to unpack social mechanisms shaping authorship, only very few contributions have attended to the role of geography in authorship attributions (see Jabbehdari and Walsh, 2017 for an exception and discussion on local versus remote collaborations). This is surprising given the rise of large geographically dispersed research projects and a more general interest in how locations and geographic proximity between researchers affect scientific knowledge production and diffusion. Studies on the geography of science have examined for instance the spatial concentration of scientific knowledge production in cities (Boschma et al., 2014; Jones et al., 2008; Nomaler et al., 2014), the extent to which scientific knowledge production is directed to local problems and needs (Ciarli and Ràfols, 2019; Confraria and Wang, 2020) and how geographic distance shapes scientific research collaboration (Catalini et al., 2020; Hoekman et al., 2010) and knowledge diffusion (Nomaler et al., 2013; Pan et al., 2012; Qiu et al., 2022; Wuestman et al., 2019). The spatial distribution of tasks and contributions in large geographically dispersed research teams and how such contributions provide local researchers with opportunities for authorship has however received limited attention so far.

This paper therefore aims to study how geography plays a role in authorship attribution to local researchers that are involved in large geographically dispersed research projects. We develop several hypotheses on how authorship opportunities in large research projects can systematically vary across researchers located in different cities depending on the environments in which they are spatially embedded and the projects in which they participate. We empirically examine these hypotheses by studying authorship attribution likelihood conditional on the execution of the same data collection task in a research project. Our baseline assumption is that researchers who contribute to data collection tasks should have the opportunity to participate in activities that are deemed necessary to qualify for authorship. Our analysis then reveals the extent to which this authorship opportunity structure is (un)equal across cities and the project and local conditions that enable or constrain researchers to realise this opportunity.

The empirical setting of our study are multicentre pharmaceutical clinical trials. Pharmaceutical clinical trials test the efficacy and safety of drugs on human health outcomes according to pre-specified protocols. Due to demands for larger sample sizes, patient diversity and generalizability as well as reproducibility of results, the size and geographic dispersion of clinical trials has expanded considerably over the last decades (Haeussler and Rake, 2017; Hoekman et al., 2012; Petryna, 2009). The accompanying rise in team size and far-going division-of-labour, has made authorship attribution increasingly complex. Observers have referred to the situation as an "*authorship lottery*" where conventions for authorship attribution are based on "*largely unwritten, but widely accepted arbitrary decisions*" (Kaufmann et al., 2010, p. 782). In this specific context, authorship attribution is not only based on substantive contributions but also market positioning of products (Polidoro and Theeke, 2012; Sismondo, 2009) and strengthening of collaborations between firms and academic researchers as key opinion leaders (Cockburn and Henderson, 1998; Moynihan, 2008). Moreover, authorship attribution comes about in unequal spatial settings with pharmaceutical trials being mainly initiated and controlled by sponsors and academic centres located in a number of clusters in high-income countries, while data collection is increasingly dispersed across the globe, including growing involvement of researchers from low and middle-income countries (LMICs) (Haeussler and Rake, 2017; Hoekman et al., 2012; Thiers et al., 2008). Geographic dispersion has raised concerns over equitable authorship opportunities and the extent to which knowledge and expertise from clinical researchers on the ground, including those in LMICs, are represented in trial authorship teams (Fisher, 2008; Kelaher et al., 2016; Obasi et al., 2021; Petryna, 2009). In a previous study Hoekman et al. (2012) showed that considerable differences in authorship opportunities exist across countries. Here we study variation in authorship opportunities across cities and add theory on geography of authorship that considers how project and local conditions enable or constrain researchers to realise authorship opportunities.

In the following, we first discuss authorship and contributorship in the context of the clinical trial research setting. We then introduce our geography of authorship perspective in Section 3 and present hypotheses to study how geography plays a role in shaping opportunities for authorship attribution. Sections 4 and 5 present the data collection and analysis strategy as well as the research findings. We end with Section 6 that positions the findings in the team science and geography of science literature and discusses various ways through which authorship policies and initiatives could foster equitable authorship opportunities in large geographically dispersed team science, independent of the location of researchers and as a fundamental principle for the conduct of science.

2. Authorship and contributorship

Authorship on publications is a main scientific institution to establish credit and responsibility for conducted research work (Biagioli and Galison, 2003; Wager, 2009). Authorship serves as a 'currency' in the reward structure of science and is instrumental for peer recognition and the accumulation of credibility (Latour and Woolgar, 1979; Merton, 1973; Stephan, 2012). Authorship also establishes responsibility and accountability for scientific work which is important in case scientific misconduct or errors are detected (Biagioli, 1998). Given these key functions, Hauessler and Sauermann (2013, p. 689) argue that "*a weak link between contributions and authorship can undermine incentives for scientific knowledge production* (Lane, 2010; Rennie et al., 1997) *as well as the scientific community's ability to enforce its norms and quality standards* (Zuckerman, 1968)".

Authorship guidelines, such as those of the International Committee

[1] We use the term authorship attribution throughout the manuscript. Attribution can follow from hierarchical decision-making in a project team as well as negotiated decision-making in self-organizing teams (Lissoni et al., 2013; Wang and Hicks, 2015). The term also applies if authors are added or left out as authors for strategic reasons, irrespective of their contributions.

of Medical Journal Editors (ICMJE)[2], The World Association of Medical Editors (WAME)[3] and other institutions[4] provide guidance regarding which contributions qualify for authorship. Despite some differences among the guidelines, a common principle is that authors can only be individuals who made substantial intellectual contributions to a study. However, even though some widely adopted guidelines – such as ICJME – provide further details on the criteria for authorship, the interpretations of which individual research tasks or combination of tasks are substantial intellectual contributions differ among scientific disciplines and research cultures.

Typically, authorship guidelines focus on individual contributions and do not take the specificities of large-scale collaborative research projects into account. This is particularly the case for those projects that spread across geographies and are characterized by considerable differences among collaborators in terms of experience, resources, research cultures and power structures. In this context, the focus on manuscript writing in some guidelines – such as ICJME – has been criticized for contributing to the potential exclusion of researchers who contributed predominantly through data collection or other technical tasks or those who have limited academic English language abilities (Penders, 2016; Smith et al., 2014). These issues have received particular attention in collaborations that involve researchers from both high-income countries and low- and middle-income countries. They also apply to a broader set of large interdisciplinary research projects that rely on the contributions of various specialized experts, not all of which will be involved in writing the manuscript (Aliukonis et al., 2020; Sauermann and Stephan, 2013).

Against this background, changes of authorship guidelines have been proposed. One set of proposals seeks to provide clarity regarding each author's contributions and to ensure that researchers with various types of contributions have opportunities for authorship in research projects that rely on profound division-of-labor (Allen et al., 2014; Holcombe, 2019; McNutt et al., 2018; Rennie et al., 1997).[5] Another set of proposed changes starts from scientific principles and norms relating to equitable participation in team science and contributing to research capacity building in the geographies where the research is conducted. Proposals from this perspective emphasize power asymmetries between researchers from different geographies. They consider providing authorship opportunities to all researchers involved a fundamental principle of science and in line with authorship guidelines[6] (Morton et al., 2022; Smith, 2023). Following this line of arguments, all researchers involved should have the opportunity to participate in the review, drafting and final approval of manuscripts, while some research may be considered illegitimate if no authors are involved in the corresponding publications from geographies where the research is conducted (Obasi et al., 2021; Smith et al., 2014).[7]

The debates on authorship attribution and the role of guidelines are reflected in findings of a longer tradition of empirical research on the type of contributions in research teams that contribute to authorship attribution. Authorship conventions and models are known to vary across disciplines and research cultures, as well as between research groups and organisations operating in the same field (Haeussler and Sauermann, 2020; Jabbehdari and Walsh, 2017; Laudel, 2002; Sauermann and Haeussler, 2017; Wager, 2009). Studies have shown for instance that contributions labelled as 'technical', including data and material provision, are considered less valuable for authorship attribution than contributions labelled as 'conceptual' (Hong, 2008; Jabbehdari and Walsh, 2017; Larivière et al., 2016; Latour and Woolgar, 1979; Sauermann and Stephan, 2013). However, the relative importance and recognition of 'technical' contributions differs per field which influences authorship attribution processes (Jabbehdari and Walsh, 2017).

Prior work also suggests that authorship may not always reflect substantive contributions. For instance, female contributors have been found to be less likely than male contributors to be attributed with authorship (Ross et al., 2022), while, focusing on lead authors, similar patterns have been reported for underrepresented racial and ethnic groups (Marschke et al., 2018). Vice-versa, authorship may not always reflect substantive contributions. Cases of honorary and ghost authorship have received considerable attention in the literature (Aliukonis et al., 2020; Flanagin et al., 1998; Gasparyan et al., 2013; Marušić et al., 2011). A survey of high-impact clinical journals revealed an average prevalence of honorary and ghost authorship of 21 % (Wislar et al., 2011). Sauermann and Haeussler (2017) report, based on an analysis of >12,000 PLoS ONE articles, that almost half of all authors do not adhere to the ICMJE criteria for authorship, primarily because of a lack of involvement in writing the manuscript (Sauermann and Haeussler, 2017). Ghost and guest authorships have also received considerable attention in relation to studies sponsored by companies. Companies may gain commercial value from including "key opinion leaders" as authors in publications, irrespective of their contribution (Moynihan, 2008) or when publications are written by professional writers who are not included in the authorship byline (Sismondo, 2009).

In the context of our empirical research setting, which is multicentre clinical trials, the typical authorship challenge relates to the contributions of local site investigators (Rosenberg et al., 2015). Local site investigators and their research teams[8] "*identify and recruit patients, conduct study procedures, complete necessary study documentation/reporting, and retain patients for outcomes assessments*" (Mentz and Peterson, 2017, p. 1185). The exact tasks and responsibilities of local site

[2] https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.

[3] https://wame.org/authorship.

[4] In addition to ICMJE and WAME, examples of institutions publishing guidance on authorship include the Council of Science Editors, the National Institutes of Health, the Committee on Publication Ethics, and The European Code of Conduct for Research Integrity. Typically, the guidelines vary in their level of detail and in the breadth of application but share some common principles, such as the idea that authors need to make a substantial intellectual contribution to the research.

[5] Several models have been proposed to acknowledge substantial contributions that do not qualify for authorship according to the ICMJE criteria. Initially, these models were based on a distinction between a role as author and contributor mentioned in the by-line or acknowledgement section on the publication (Hawkins, 2020). The development of such models followed from editors calls to not 'simply' list all investigators from large research projects such as multicentre clinical trials as authors on a publication (Kassirer and Angell, 1991; Rennie et al., 1997). The so-called Rennie-Yank-Emanual system acknowledges all investigators in credit rosters published along the manuscript and includes a description of the nature of their contribution. This system was initially adopted by a number of medical journals (Yank and Rennie, 1999) and later led to more widespread uptake of contribution disclosures in journals based on role taxonomies (Sauermann and Haeussler, 2017). One of those taxonomies, the Contributor Roles Taxonomy (CRediT) was endorsed by the authorship guideline recommendations of McNutt et al. (2018) which were then adopted by many publishers including Nature, BMJ and Cell Press. It is envisioned by some that the uptake and visibility of these contribution disclosures will increase to such an extent that it may ultimately provide a substitute for (order of) authorship altogether.

[6] For instance, the ICMJE guideline states that "criteria are not intended for use as a means to disqualify colleagues from authorship who otherwise meet authorship criteria by denying them the opportunity to meet criterion #2 or 3. Therefore, all individuals who meet the first criterion should have the opportunity to participate in the review, drafting, and final approval of the manuscript."

[7] Illegitimate research has particularly been linked to practices that have become known under the label of parachute research, helicopter research or parasitic research. Morton et al., 2022, p. 265 defines those practices as the conduct of primary research within a host country and subsequently publishing findings with inadequate recognition of local researchers, staff and/or supporting infrastructure (Morton et al., 2022).

[8] We simply refer in this paper to local site investigators.

investigators may differ depending on project characteristics such as the phase of the clinical trial, the trial sponsor, the negotiated division of labour within a team of investigators, and other factors. For instance, early-stage trials (phase 1 and phase 2) focus predominantly on the generation of new knowledge through the development and the testing of hypotheses (Azoulay, 2004). In contrast, phase 3 trials focus predominantly on the generation of empirical evidence and not so much on hypothesis development. These trials require the involvement of a large number of local site investigators who are often distributed across geographies (Haeussler and Rake, 2017).

However, as particularly phase 3 – but also some phase 2 clinical trials - require the involvement of a large number of local site investigators, it can be difficult to include them all in the manuscript writing process to fulfil authorship guideline criteria (Rosenberg et al., 2015). In these cases, "*little guidance exists surrounding authorship attribution processes when the number of researchers in a trial exceeds that which can be negotiated by discussion and consensus alone*." (Whellan et al., 2015, p. 458). Often, principal investigators organised in clinical trial management teams are involved in conceptualizing and designing the research as well as interpreting results, while local site investigators are typically predominantly involved in data collection. Local investigators might be offered opportunities to contribute to the manuscript writing processes and/or to become authors on a resulting paper (Hoekman et al., 2012). They may however have different motivations and incentives to become authors. Some investigators are motivated by opportunities to advance the state of knowledge or to advance clinical practice as well as by opportunities to get academic credit for their contributions through publications. Others may be predominantly motivated by monetary compensations or by other factors and may have less interest in authoring academic publications (Rasmussen, 2005; Rettig, 2000).

As common authorship guidelines provide little guidance on how authorship should be determined in these settings, practices to attribute authorship to local site investigators tend to vary between clinical trials. Authorship agreements may be specified in an authorship contract at early stages of the research process or determined later in the process based on actual contributions or other decision heuristics. Moreover, authorship attribution might be based on formalized authorship attribution methods that aim to make the decision on which investigators become authors more transparent and equitable (Dulhunty et al., 2011; Whellan et al., 2009).[9] In contrast, they might also follow from local routines or idiosyncratic practices proposed by clinical trial leadership and then accepted by local site investigators who participate in the trial (Archer et al., 2016; Hawkins, 2020). In large trials, clinical trial management teams, constituted in interaction between the principal investigator and sponsor, often become involved in manuscript writing and are attributed authorship. Management teams might also decide to attribute authorship to local investigative researchers who might contribute to the manuscript writing process or be credited for other activities.[10]

Based on this we conclude that authorship attribution in clinical trials thus reflects a diverse mix of conventions and practices that can be established during various phases of the research process. Below, we argue that such practices and conventions can contribute to systematic differences in researchers' authorship opportunities across locations, to convert their contributions into authorship attribution. Our baseline assumption is that researchers who contribute to data collection tasks should have the opportunity to participate in activities that are deemed necessary to qualify for authorship. Our hypotheses regarding the geography of authorship are then meant to examine this authorship opportunity structure from a geographic perspective by considering how authorship opportunities for local researchers vary depending on the environments in which local researchers are spatially embedded and the projects in which they participate.

3. Geography of authorship

Facilitated by decreased costs of travelling and advances in communication and information technologies, tasks in large research projects have become increasingly geographically dispersed, with single places being involved in multiple tasks, and the same task being distributed over multiple places (Catalini et al., 2020; Hoekman et al., 2010). Geographic dispersion can be expressed in terms of divisions-of-labour between places, task standardization across places and hierarchical control from specific places (Walsh and Lee, 2015). Large research projects, and in particular interdisciplinary ones, are known to rely on specialized resources, infrastructure and expertise sometimes only available in specific places (Haeussler and Sauermann, 2020). Tasks in these projects are often standardized to facilitate synchronous execution in different places and centralised coordination from scientific command-and-control centres.

Geography of authorship considers how authorship opportunities and the conversion of contributions into authorship attribution vary across geography among researchers in large, bureaucratically organised research projects. We argue that opportunities for local investigative researchers to become authors in large research projects vary depending on how local researchers are embedded in projects and local environments. Importantly, these mechanisms shaping authorship attribution opportunities can be conceptualised relatively independently from mechanisms shaping the spatial distribution of contributions in the first place. Thus, while the geography of scientific contributions, including the geography of clinical trials, is highly uneven across space (Boschma et al., 2014), there are no a-priori reasons to assume that opportunities for researchers to become authors in large research projects reflect this unequal distribution, in case researchers make similar contributions in projects.

[9] For instance, authors from the HF-ACTION trial conducted at 82 sites in 3 countries developed and used a score-based system to rank the performance of clinical trial sites. Site-specific performance metrics are generated based on the entire trial process from patient enrolment, adherence to the intervention, data completion, and investigators' participation in trial committees and oversight of laboratory operations (Whellan et al., 2015, 2009). Generated scores are combined with the preferences of site investigators for writing a particular type of manuscript. In case a site has a designated author on a publication, points are deducted in order to maintain an equal distribution of authorship among sites across publications (Whellan et al., 2015, 2009).

[10] Formalized authorship attribution methods have been developed that aim to make the decision on which investigators become authors more transparent and relatively equitable, also with the goals of stimulating exchange of knowledge between investigators and maximise knowledge dissemination (Dulhunty et al., 2011; Whellan et al., 2015). For instance, authors from the HF-ACTION trial conducted at 82 sites in 3 countries developed and used a score-based system to rank the performance of clinical trial sites. Site-specific performance metrics are generated based on the entire trial process from patient enrolment, adherence to the intervention, data completion, and investigators' participation in trial committees and oversight of laboratory operations (Whellan et al., 2015, 2009). Generated scores are combined with the preferences of site investigators for writing a particular type of manuscript. In case a site has a designated author on a publication, points are deducted in order to maintain an equal distribution of authorship among sites (Whellan et al., 2015, 2009).

3.1. Embeddedness in project

Starting with embeddedness in large research projects, it can be expected that opportunities for local investigative researchers operating from a particular location to convert contributions into authorship attribution are dependent on the sheer number of locations involved in a project. This is in line with individual-level evidence on authorship attribution which shows that team size is positively associated with functional specialization and, particularly, with a high level of task division in empirical activities such as data collection (Haeussler and Sauermann, 2020; Walsh and Lee, 2015). When research projects involve more locations for data collection, these locations are thus more likely to be specialized and less likely to be involved in other tasks that would increase opportunities for authorship attribution. When the sheer number of involved locations becomes larger, locations also experience competition for authorship, which is particularly strong when group authorships are used which provide credit to local investigative researchers instead of listing all local investigators on the by-line (Rennie et al., 1997).

Given the number of locations involved, authorship attribution opportunities are also expected to differ depending on the extent to which tasks are standardized and hierarchically controlled. Variation in standardization and hierarchical control have a spatial expression as they direct the flow of information, data and knowledge between involved locations as well as the extent to which geographic proximity is important for knowledge transfer between local investigative researchers. In general, we expect that both higher levels of standardization and stricter hierarchical control reduce authorship opportunities for local investigative researchers. This is because standardization of empirical activities (e.g., through data collection protocols and procedures) reduces the need for face-to-face interactions between locations and facilitates vertical organisation of information, data and knowledge transfer over horizontal ones. Standardization and hierarchical control thus increase possibilities for distant involvement of locations and decoupling of more 'technical' information-intensive processes from 'conceptual' knowledge-intensive processes such as research design, interpretation of data and manuscript writing (Azoulay, 2004). As a result, when standardization and hierarchical control are substantial, authorship networks can be organised in a relatively closed fashion (Wang and Hicks, 2015).

In the specific context of clinical trials, it is known that information and knowledge flows are more strictly controlled and vertically organised in industry-sponsored studies compared to publicly-sponsored and investigator-initiated studies. Due to the commercial interests of pharmaceutical companies in the outcome of studies, hierarchical control is particularly exercised over the disclosure of research findings and access to key resources (e.g., proprietary data, infrastructure) that enable contributions to disclosure processes. In line with commercial logics in science (Dasgupta and David, 1994; Sauermann and Stephan, 2013), investigators in industry-sponsored trials might be hired as contract researchers with "*little influence over what types of studies are done, what questions are asked, what procedures required, and which patients can qualify*" (Fisher, 2008, p. 35). Local investigative researchers in industry-funded trials also sometimes report that they might not be aware of how authorship decisions are made and lack access to data sources limiting their opportunities to make contributions to data analysis and interpretation (Rasmussen et al., 2018; Tauber and Paul, 2017). Moreover, their relationships with sponsors might be arm's length and mediated by clinical research organisations (CROs) that act as subcontractors of the sponsor to recruit, coordinate and supervise data collection at specific locations (Azoulay et al., 2010; Fisher, 2008). We thus formulate the following hypothesis:

Hypothesis 1. The likelihood of authorship attribution to local researchers in multi-city research projects is lower when they contribute to industry-sponsored projects compared to non-industry sponsored projects.

Project embeddedness also refers to the geographic location of local investigative researchers vis-à-vis other involved researchers and the centres from which research projects are coordinated. Studies on the geography of research collaborations have shown that research collaborations continue to be sensitive to geographic distance, despite advances in information and communication technologies (Catalini et al., 2020; Hoekman et al., 2010). The need for copresence and more sustained forms of colocation between researchers is particularly high for the exchange of complex information and tacit knowledge (Collins, 2010; Polanyi, 1966). Geographic proximity is also more important when cognitive distances need to be bridged such as in interdisciplinary research teams and in knowledge exchanges between organisations that work under different institutional regime structures (Ponds et al., 2007).

For the specific context of clinical trials, coordination and management tasks of clinical trials can be organised in close proximity to local researchers working in a clinical environment in order to facilitate knowledge exchange between science and clinical practice (Gittelman, 2016). Geographic proximity facilitates mechanisms of local knowledge exchange between scientists and clinicians such as informal communication, joint clinical observations, and serendipitous encounters (Gelijns and Rosenberg, 1994). This contributes to a natural coupling of more technical data-intensive and conceptual knowledge-intensive tasks. Such spatial coupling of conceptual and data-intensive tasks is expected to be more likely when local researchers are relatively close to leadership and coordinating centres, translating into increased authorship attribution opportunities:

Hypothesis 2. The likelihood of authorship attribution to local researchers in multi-city research projects is positively related with the geographic proximity of the local researcher to the coordinating sponsor.

Local researchers also face competition for authorship. Besides project-level competition which reduces the likelihood of authorship for all local teams involved, competition also has a spatial component in large geographically dispersed research teams. We expect that such spatial competition mainly operates within rather than across countries based on three different mechanisms. First, the spatial organisation of clinical trials can be hierarchical, with national or regional coordination centres being responsible for data management, monitoring and facilitating knowledge exchange between different local teams in a country or region (Petryna, 2009). When such internal organisational structures apply, it is likely that only one coordinating centre in a country or region provides authors on a publication, given that more knowledge intensive tasks are executed at those centres. Second, the credibility of multicentre clinical trials might increase when trials are conducted in multiple countries due to the increasing generalizability and external validity of findings. One way to signal this diversity and the associated quality of the trial is through diverse authorship attribution to countries. As such local teams compete for authorship within countries, but less so between countries. Third, clinical trials are increasingly conducted in non-traditional research locations that are attractive as a market for pharmaceuticals (Haeussler and Rake, 2017). Authors from these countries may be selected as "key opinion leaders" that can contribute to the diffusion of knowledge on the experimental treatment and its potential benefits in national clinical networks and markets (Moynihan, 2008; Sismondo, 2009). Thus, authorship might be driven by mechanisms of spatial competition and representation at the country level, leading to our third hypothesis:

Hypothesis 3. The likelihood of authorship attribution to local researchers in multi-city research projects is negatively related with the level of national competition that local researchers face.

3.2. Embeddedness in local environment

The geography of scientific knowledge production is highly uneven

(Heimeriks and Boschma, 2013). Spatial concentration of science is particularly strong in the biomedical sciences that requires alignment between a highly distributed competence base, involvement of a heterogeneous set of actors and bi-directional knowledge transfer between scientific research and clinical practice (Gelijns and Rosenberg, 1994; Mina et al., 2007). Concentration of scientific knowledge production in the biomedical field is facilitated by mechanisms of local knowledge exchange such as informal communications, serendipitous encounters, social and professional contacts and labour market mobility (Almeida and Kogut, 1999; Breschi and Lissoni, 2009; Malmberg and Maskell, 2002). It is also strongly dependent on the local presence of institutional complementarities between scientific and non-scientific organisations (e.g., hospitals and companies) and technical complementarities in terms of infrastructure and material (Bonaccorsi, 2010).

Local researchers in large multi-city research projects vary in how they are locally embedded in such environments. Local embeddedness provides them with differential means (e.g., resources, expertise, infrastructure, credibility) to execute particular tasks, make project contributions and receive credit and recognition. In general, locations differ in terms of accumulated experience with clinical trial conduct and involvement in publication processes which is materialised in infrastructure, coordinated networks and institutionalised practices. On an individual level it has been shown that such accumulated experience in the form of past research performance is important for authorship attribution, independent of team science contributions (Haeussler and Sauermann, 2013; Lissoni et al., 2013, 2020). Akin to this, we expect that experience-based variation between locations in clinical trial conduct and publication also shapes authorship opportunities.

On top of such differences in experience, we expect that reputation and status of researchers and the locations from which they contribute play a role. Authorship of high-status researchers and - by extension high-status universities and locations - may serve as a quality signal on publications enhancing the legitimacy, visibility and diffusion of research (Simcoe and Waguespack, 2010). Dominant valuation logics in the science system through which such status and reputation can be signalled are citation-based metrics (Wouters, 1997) as well as university rankings which provide a normative isomorphic framework for evaluating differences in performance and accomplishments (Frenken et al., 2017). We expect that these spatial differences matter for authorship attribution and formulate the following hypothesis:

Hypothesis 4. The likelihood of authorship attribution to local researchers in multi-city research projects is positively related with the scientific reputation of their local environment.

Knowledge produced by local researchers and the project in which they participate might be more or less aligned with their local environment. Scientific knowledge production in the medical field proceeds in so-called problem sequences that evolve through the search for increasingly specific solutions to clinical problems (Metcalfe et al., 2005; Mina et al., 2007). However, there is considerable geographic variation in the nature and burden of clinical problems as well as the specific local needs and demands in terms of solutions (Confraria and Wang, 2020). Application and deployment of solutions is also context-specific, requiring local expertise, knowledge, and capacity-building to serve needs. This perspective raises attention to the alignment of research activities with local needs and research priorities (Ciarli and Ràfols, 2019; Confraria and Wang, 2020).

In the context of authorship attribution, there are several reasons to expect that local researchers operating in an environment where particular clinical problems are relatively prevalent have a higher likelihood of becoming an author on scientific publications that test interventions that address these problems. These researchers might play a larger role in such research projects due to greater possibilities to develop knowledge on the problem-solution pair at hand. They can also bring in necessary local expertise and knowledge to tailor the development of interventions to contexts where the problem is most severe. Finally, local researchers can champion solutions in their local environment and act as key-opinion leaders for local diffusion of research findings (Moynihan, 2008; Sismondo, 2009). This leads us to our fifth hypothesis:

Hypothesis 5. The likelihood of authorship attribution to local researchers in multi-city research projects is positively related with the extent to which the research project addresses a local problem.

4. Methods

To study authorship attribution in multi-city clinical trials, we created a dataset of clinical trials that were conducted in at least two different cities. The study entailed an extensive data collection effort linking registered clinical trials in the U.S. National Library of Medicine (NLM) web-based registry ClinicalTrials.gov with corresponding publications indexed in the NLM PubMed database and the bibliographic database Scopus Elsevier. In a first step, we extracted city information of participating facilities in registered clinical trials and author address information from corresponding publications. In a second step, we allocated trial facilities and publishing cities to urban areas across the globe. In a third step, we determined for each trial whether an author from an urban area was an author on a trial-related publication conditional on the presence of at least one participating facility from that urban area in the clinical trial. In a final step, we collected additional data on urban areas and registered trials in line with the formulated hypothesis. In the following, we explain the process of constructing our dataset and linking different data sources in more detail.

4.1. Allocation of facilities and author addresses to urban areas

We obtained the full set of information on clinical trials registered in ClinicalTrials.gov by downloading the Clinical Trials Transformation Initiative's database for aggregate analysis of ClinicalTrials.gov on May 24, 2018.[11] We then linked registered trials to PubMed indexed publications based on an exhaustive search of ClinicalTrials.gov registry numbers in PubMed's secondary source ID field. In case publications were linked to a registered clinical trial, bibliographic data from the publication including author addresses, affiliations and positions were retrieved from Scopus Elsevier.[12]

All participating facilities and author addresses were subsequently allocated to urban areas, using a self-constructed spatial database that covers 1875 of the largest urban areas in the world (see Appendix 1). The database follows the functional urban area (FUA) definition of the Organisation for Economic Co-operation and Development (OECD). OECD defines urban areas as densely populated urban centres with at least 50,000 inhabitants and the commuting zone of these centres. It covers cities in 38 OECD or European Union (EU) countries (Dijkstra et al., 2019). For other non-EU or non-OECD countries, we included urban areas of cities with >500,000 inhabitants using population grids from global human settlement databases (see Appendix 1 for sources).

To obtain geographic coordinates of all trial facilities and author addresses and be able to allocate them to urban areas, we used the geocoding webservice of GeoNames.[13] Retrieval of coordinates of facilities and author addresses was done based on combined queries of city, state (only United States), and country names. Geospatial techniques were then used to assess whether a facility or author address was located within the boundary of an urban area. Appendix 1 provides more detail on urban area definitions and the results of the geocoding process. In the following we simply refer to cities instead of urban areas.

[11] https://ctti-clinicaltrials.org/.

[12] Scopus Elsevier is a comprehensive bibliometric database that covers >23,000 journals with an extensive coverage of health and life sciences.

[13] https://www.geonames.org/.

4.2. Sample and dependent variable

In a second step we selected a sample of clinical trials to construct an analysis dataset. Based on our sample definition we included clinical trials registered at ClinicalTrials.gov if they were completed in the period 2008–2017, tested at least one drug or biological intervention and were multi-city clinical trials defined as having an enrolment facility in at least two cities. In terms of the link of trials to corresponding publications, we included clinical trials if they were linked to multiple publications because the results of large clinical trials can result in multiple publications that might be part of prospective publication plans (Hawkins, 2020). We excluded publications that were linked to multiple clinical trial registrations as such publications often report on data of pooled-analysis or meta-analysis performed by other research teams than those conducting the trial. Appendix 2 provides a flowchart of the sample construction process.

We then constructed a binary dependent variable capturing authorship attribution to a city conditional on the participation of that city with at least one facility in the clinical trial. Thus, our dependent variable takes on a value of one in case a city with at least one trial facility is present on one of the corresponding publications, and zero in case a city with at least one trial facility is not present on one of the corresponding publications.

4.3. Independent variables

In line with our hypotheses the registered trials and cities of trial facilities in the sample were subsequently characterized based on whether they are industry-sponsored (H1), distance to coordinating sponsor (H2), national competition (H3), scientific reputation (H4) and local problem (H5).

Industry-sponsored (*H1*): For each registered trial, we extracted information on the lead sponsor from Clinicaltrials.gov. The dataset already classifies sponsor types in Industry, NIH, U.S. Fed and Other types. We recoded sponsor type into a variable coded 1 in case of an industry-sponsored trial and 0 otherwise. The latter category mainly consists of publicly-funded trials by research agencies and councils but also includes some trials funded by universities or NGOs.

Distance to sponsor (*H2*): For all sponsors we manually determined the city location of their global headquarters based on information from sponsors' websites, Crunchbase[14] and media sources. Clinicaltrials.gov distinguishes between lead sponsors and collaborators and we focused on lead sponsors. We extracted the city of the global headquarters of the lead sponsor, even for rare case where the sponsors' name made explicit reference to subsidiaries or research facilities in specific locations. In case an individual was mentioned as a sponsor, the city location of the affiliated institute was used. In case a sponsor did not have a headquarters (e.g., in case of academic collaborative networks) the location of the main coordinating organisation was considered the headquarters. The geographic coordinates of all headquarters were obtained using similar geocoding procedures as described above. Geographic distance to the sponsor's headquarters was computed as the geodesic distance between the headquarters city and the centroid of the facilities' city.

National competition (*H3*): We constructed a simple measure of national competition defined as the proportion of investigative sites in a clinical trial that are located in the same country but not in the same city. A high proportion thus indicates that a city competes for authorship on the national level with a relatively high number of other cities in that country.

Scientific reputation (*H4*): We constructed citation-based measures obtained from the Leiden University Ranking[15] to determine the scientific reputation of universities in the city in which facilities are located. The Leiden University Ranking 2020 contains citation-based indicators for a large set of 1176 universities across the globe (Waltman et al., 2012). We determined the main city location of each university, obtained geographic coordinates for locations and allocated locations to the cities in our sample. We then used a specific impact measure for the biomedical and health sciences field expressed as the proportion of university-produced publications in area city that belong to the top 10 % most frequently cited publications in that field. The data is available in four-year periods starting with the period 2006–2009. Scientific reputation was computed for the four years prior to the completion year of the clinical trial. For the years 2008 and 2009, data for the period 2006–2009 was used.

Local problem (*H5*): To determine the severity of the local clinical problem associated with the indication being tested and the need for the tested intervention in the trial we relied on the Global Burden of Disease (GBD) studies of the World Health Organisation (Murray et al., 1996; WHO, 2018). These studies provide country-level data on mortality and disability from diseases, injuries, and risk factors. This includes a single measure of disability adjusted life years (DALYs) to quantify disease burden for a specific disease (Murray et al., 1996). DALY measures disease burden based on the time lost due to premature death (referred to as mortality) as well as the time lived in a less than optimal state (referred to as disability). One DALY represents one lost year of 'healthy' life, and the measured disease burden is the gap between a population's health status and that of a global reference population (WHO, 2018).

Previous studies have used DALY measures to assess research funding as well as publication activities against disease burden (Confraria and Wang, 2020; Marshall et al., 2021). We use similar methods as documented in Marshall et al. (2021) to link registered clinical trials to level 3 cause categories of disease burden (see Appendix 3 for a list). To do so, all terms listed in the conditions and condition_browse field of ClinicalTrials.gov were linked to Medical Subject Headings (MeSH) terms. The terms mentioned in these field represent the primary disease or condition being studied in the trial and are expected to be entered in the registration using the MeSH controlled vocabulary. We then linked these MeSH terms to ICD-10 codes terms using the Unified Medical Language System (UMLS) that provides synonyms between different biomedical and health vocabularies including synonymous MeSH and ICD-10 terms. ICD-10 codes were subsequently mapped onto level 3 GBD cause categories using the concordance table available from the Global Disease Burden Study (WHO, 2018).

We relied on country-level GBD estimates for the year 2016 and retrieved for each facility corresponding DALYs per 100,000 inhabitants for the disease(s) being studied. In case multiple diseases were mentioned for a single trial, average DALY values were taken. Moreover, in this step we also excluded a number of trials because their disease focus could not be linked to corresponding DALY estimates (see flowchart in Appendix 2).

4.4. Control variables

Several variables known or expected to be associated with authorship attribution are included as control variables in the analysis. We grouped the control variables in terms of whether they pertain to the clinical-trial level or city-level and exclude clinical-trial level controls in our fixed-effects models (see below).

Trial-level control: We expect that the likelihood of authorship attribution might vary depending on the size as well as complexity of the trial. We therefore include measures of the total number of enrolled patients (*total patients*) and the total number of facilities in the trials (*total sites*) as well as the trial duration in days from first patient enrolment to date of last data collection (*duration*). We expect that the inclusion of more sites reduces authorship likelihood, while higher numbers of patient enrolment as well as a longer duration of trials increases authorship rates as it indicates more substantial trial involvement. We also include a number of other trial characteristics that are

[14] www.crunchbase.com.
[15] www.leidenranking.com.

known to contribute to variation in clinical trial standards, design and execution. As Phase 3 trials are generally larger, more data- and cost-intensive and follow strictly pre-determined protocols (Azoulay, 2004; Haeussler and Rake, 2017) we include a dummy variable for Phase 3 trials (*Phase 3*). Trials can also differ in complexity depending on the intervention being tested. We therefore include a dummy variable for those clinical trials that test only small molecule drugs versus trials that also test other more complex interventions such as biologicals or non-drug interventions (*Small molecule*). Finally, it is known that standards for and design of clinical trials differ considerably between disease areas. We therefore add disease dummies based on the level 2 GBD cause categories to our models.

City-level controls: Authorship attributions might be positively associated with experience with conducting trials as well as writing scientific publications. We therefore included for each city a measure of the log transformed number of trials (*trial experience*) registered in Clinicaltrials.gov as well as log-transformed number of university-produced publications based on the Leiden University Ranking (*publication experience*) in respectively three and four years prior to the completion of the clinical trial. We also expect that the likelihood of authorship attribution to a city increases with the number of facilities from that city in the clinical trial and thus include the log-transformed number of city facilities in the trial (*city sites*) as a control variable. Finally, prior studies have shown that sponsors are likely to appear as co-authors on clinical trial publications, particularly in case of industry-sponsored trials (Buchkowsky and Jewesson, 2004; Rafols et al., 2013). Given our interest in authorship by participating facilities we control for this by including a binary variable taken on a value of one in case the city is also the location of the study's sponsor headquarter, with an expected positive effect on authorship attribution.

4.5. Data analysis

We report descriptive analyses and estimate regression models. Our dependent variable for the regression models is binary and follows a binomial distribution. We therefore use logistic regression techniques to predict authorship likelihood of cities. As the city observations are nested in clinical trials, we estimate conditional logit models and include a trial-level fixed effect that controls for unobserved heterogeneity at the trial-level. Moreover, as cities occur repeatedly in our dataset, we cluster standard errors in our models across cities. Data was analysed using R and the fixed-effects logistic models specifically using the survival package and *clogit* command using the Efron method for maximum likelihood estimation.

5. Results

5.1. Clinical trials

Our sample consists of 3619 clinical trials that are linked to at least one publication. The majority of these clinical trials are industry-sponsored (71.0 %), international (55.7 %), test only small molecule drugs (67.2 %) and are either phase 3 trials (43.2 %) or phase 2 trials (33.5 %). These trials enrol a median of 217 patients [Interquartile range (IQR): 77–518] for a median of 830 days [IQR 510–1370 days]. In terms of disease areas, we observe relatively high proportions of trials for malignant neoplasms (20.8 %) and infectious and parasitic diseases (14.2 %) (see Fig. A3.1 in Appendix 3 for complete overview). We also observe that industry-sponsored trials tend to enrol more patients and are more often international and phase 3 trials, yet shorter in duration. Phase 3 trials also tend to enrol more patients and are more often international, yet equal to other trials in terms of duration.

Patient enrolment in these clinical trials is geographically dispersed with a clinical trial having at least one investigative site in a median of 15 [IQR: 5–41] and average of 31.1 [Standard Deviation (SD): 43.9] cities. This amounts to a total number of 112,503 city observations in our dataset. Fig. 1 shows the geographic distribution of clinical trials over cities. We observe that clinical trial conduct is mainly concentrated in cities in North America and Europe which together constitute 85.9 % of all city observations in the dataset, while conduct in cities in Africa (0.9 %), Oceania (2.1 %) and South America (3.1 %) is relatively low.

The 10 cities with the highest number of trials are all located in the United States, while London, Berlin, Paris, and Toronto are non-US cities with relatively high number of clinical trials. When distinguishing between traditional and emerging regions for clinical trial conduct (Haeussler and Rake, 2017; Thiers et al., 2008),[16] we observe that 24.0 % of city observations are in non-traditional regions. When further distinguishing between income groups according to the World Bank country classification by income, we observe that 86.2 % of cities are located in high-income countries, 10.7 % in upper-middle income countries, 3.1 % in lower-middle income countries and 0.09 % in low-income countries. At the clinical trial level, 42.7 % of trials have at least one recruiting facility in a non-traditional region for clinical trial conduct, while 34.3 % of trials have at least one recruiting facility in a LMIC.[17]

5.2. Publications and authorship

The 3619 clinical trials are linked to a total number of 5376 publications, with 78.0 % of trials being linked to only one publication. We find that these publications list on average 9.86 authors. On an average publication, 66.1 % of authors are located in cities with at least one investigative site in the trial, 27.5 % in cities with no investigative sites in the trial (e.g., sponsor affiliations) and 6.4 % of authors are in locations outside of the cities in our dataset.

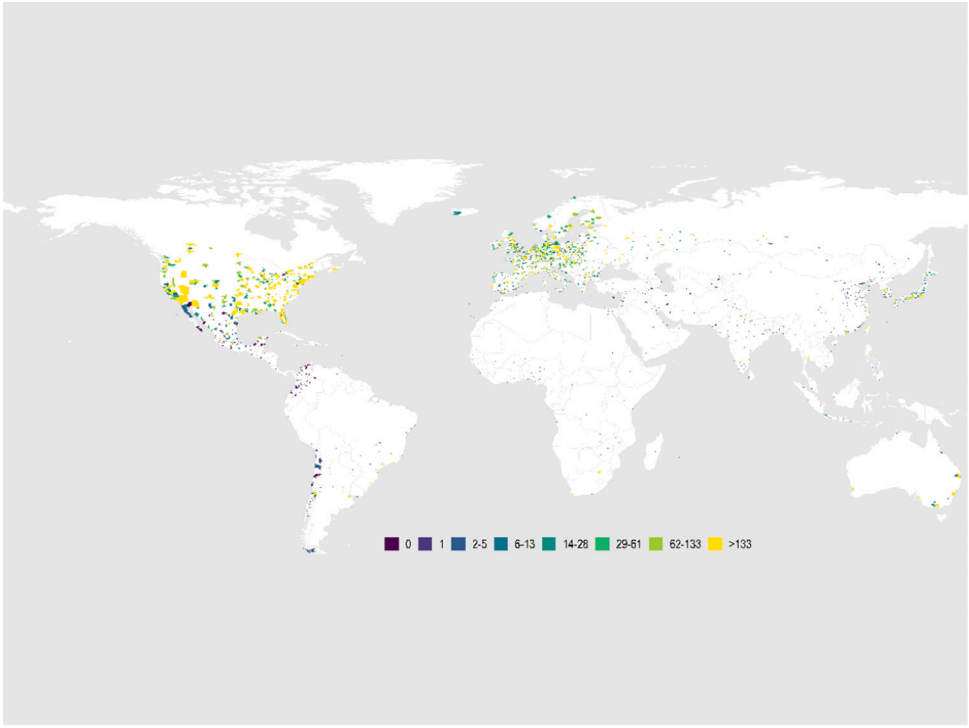
In terms of authorship, 16.7 % of city observations have at least one author on a corresponding publication. Fig. 1b shows proportions of authorship attribution by city. Authorship attribution occurs almost twice as often for cities in traditional locations compared to emerging research locations (19.0 % versus 9.9 %). It is also higher for high-income countries (17.9 %) compared to upper-middle income (9.9 %) and lower-middle income countries (7.7 %), and particularly high for low-income countries (82.5 %). When considering authorship rates per city, we observe a very weak positive correlation between authorship rates and the number of clinical trials in a city ($r = 0.19$) when considering all cities and a moderate positive correlation ($r = 0.46$) when only considering cities with at least 10 clinical trials. A complete overview of authorship rates per city and country is provided in Appendix 4.

The median (IQR) percentage of cities that are attributed authorship per clinical trial is 28.6 % (IQR = 9.3 % - 83.3 %). This percentage however decreases sharply with the total number of cities involved in a trial. Fig. 2 shows boxplots of authorship percentages for cities stratified by the total number of cities involved in a clinical trial divided in octiles. While the median number of cities being attributed authorship is 100 % in the lower 25 % of the distribution ($\leq$5 cities), these percentages decrease sharply when the number of cities increase. Authorship percentages are below 25 % above median size ($\geq$16 cities) and <10 % for the upper quartile ($\geq$42 cities). Fig. 2b further breaks down these authorship proportion by whether trials are industry-sponsored or not. The Figure shows that, given a particular clinical trial size, opportunities for authorship attribution are lower for cities participating in industry-sponsored trials compared to non-industry sponsored trials.

[16] Traditional regions are defined in Haeussler and Rake (2017) as United States, Canada, Japan, Australia, New Zealand and the 15 EU member states prior to the accession of ten Eastern European candidate countries in 2004.

[17] These observations are comparable to previously reported numbers in the literature (e.g., Awan et al., 2022; Haeussler and Rake, 2017; Thiers et al., 2008).

a. Number of clinical trials per city



b. Authorship proportionper city

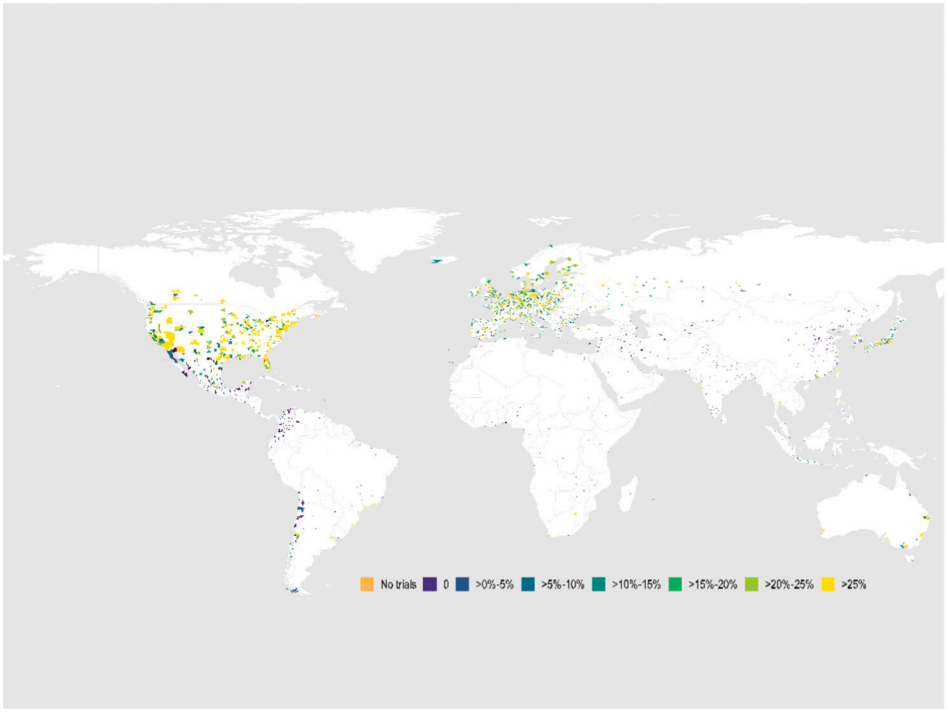
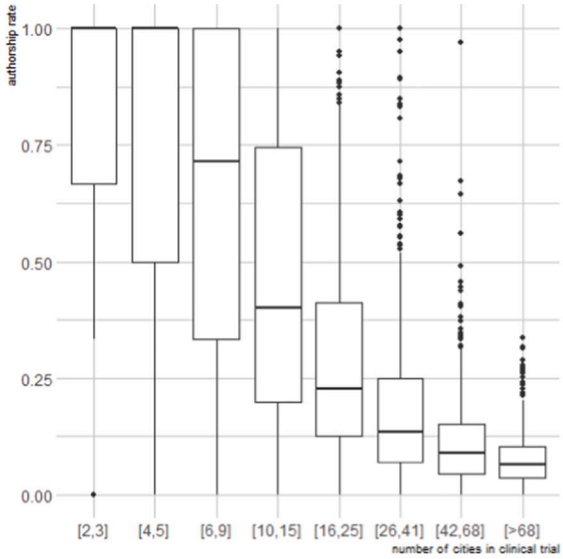


Fig. 1. a. Number of clinical trials per city.
b. Authorship proportion per city.
Note: raw data is provided in Appendix 4.

a. Authorship proportionfor cities by total number of cities in clinical trials



b. Authorship proportion for cities by total number of cities and funding source

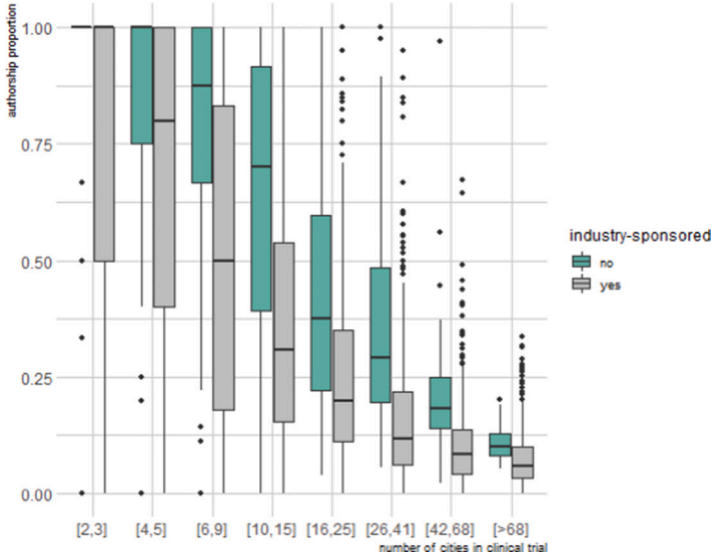


Fig. 2. a. Authorship proportion for cities by total number of cities in clinical trials.
b. Authorship proportion for cities by total number of cities and funding source.

5.3. Regression results

Table 1 lists descriptive statistics and Table 2 the correlation matrix for the variables and city observations included in the regression models. Correlations between independent variables are generally weak, except for relatively strong positive correlations between total patients and total sites ($r = 0.80$), scientific reputation and trial experience ($r = 0.65$), scientific reputation and publication experience ($r = 0.86$) and publication experience and trial experience ($r = 0.68$).

Tables 3 and 4 present the estimates from the logistic regression models on authorship likelihood of cities. Table 3 adds the trial-level and urban-area level control variables (Model 1 and 2) and tests hypotheses 1 (Model 3) which is formulated on the project-level and predicts that overall authorship likelihood of cities is lower in industry-sponsored projects compared to other projects. The coefficient for this covariate is negative as expected and statistically significant ($\beta = -0.625, p < 0.001$).

Looking at the trial-level control variables we observe a strong negative effect of the total number of sites in the trial and a positive and significant effect of the total number of patients and duration of the trial. We also observe a relatively small but significant positive effect of being a Phase 3 trial on authorship likelihood. The effect of the type of intervention is insignificant.

The models presented in Table 4 add fixed effects for clinical trials and test hypotheses 2–5 formulated on the city level using the conditional logit specification. Models 2–5 add independent variables sequentially according to our hypotheses, Model 6 presents a full model.

Table 1
Summary statistics.

Statistic	Mean	St. Dev.	Min	Median	Max
Authorship (dv, b)	0.17	0.37	0	0	1
Industry-sponsored (b)	0.88	0.33	0	1	1
Distance to sponsor	4637.70	4028.57	0.56	3476.69	19,054.50
National competition	0.84	0.21	0.00	0.92	0.999
Scientific reputation	0.08	0.06	0.00	0.09	0.21
Local problem	533.79	683.71	0.46	312.83	11,963.13
Trial controls					
Total patients	2210.78	4882.62	3	599	84,496
Duration	1136.47	793.30	1	929	6605
Phase 3 (b)	0.68	0.47	0	1	1
Small molecule (b)	0.67	0.47	0	1	1
City controls					
Total sites	183.51	239.00	2	101	1621
City sites	1.49	1.64	1	1	170
Trial experience	616.88	882.95	0	272	6033
Publication experience	3654.85	5577.68	0	1545	35,117
Sponsor headquarter (b)	0.01	0.11	0	0	1

112,503 observations; (dv) indicates dependent variable; (b) indicates binary variable.

We discuss regression results based on the full model.

Starting with the control variables, we observe expected effects. Authorship likelihood of cities decreases with the number of facilities in the trial, while it increases with the number of contributing facilities in the respective city. We also find positive and significant effects for experience both with prior clinical trial involvement and publication experience. The dummy variable indicating that a city is the home location of the sponsor's headquarter shows a positive significant effect. This might point towards frequent authorship of sponsors on publications or to high authorship likelihoods for investigative centres that are located in the same city as the sponsor.

Hypothesis 2 holds that authorship likelihood of cities decreases with the distance from the sponsor's headquarter. We indeed find a negative and significant effect ($\beta = -0.066$, $p < 0.001$).

Hypothesis 3 tests the effect of national competition among cities. In line with the hypothesis, we find a negative and significant effect of national competition ($\beta = -0.545$, $p < 0.001$). Importantly, the introduction of the variable does not strongly affect the coefficient of overall competition in the project as expressed by the number of facilities involved, suggesting that spatial competition seems to operate relatively independent from overall project-level competition.

Hypothesis 4 estimates the effect of city's scientific reputation on authorship likelihood and finds the expected positive and significant effect. The effect of local excellence is relatively strong compared to the effects observed for the other hypotheses ($\beta = 3.036$, $p < 0.001$).

Finally, hypothesis 5 predicts that authorship likelihood of a city increases with the local severity of the problem addressed in the research project. We also do find a positive significant effect for this variable ($\beta = 0.144$, $p = 0.001$).

5.4. Stratification

As industry-sponsoring has a considerable effect on the overall likelihood of authorship attribution, we explore differences in regression estimates between industry-sponsored and non-industry sponsored trials. Moreover, as lead authorship attribution (i.e., first or last author) in the form of a first or last authorship is generally considered to be more reputable and visible than middle-authorship we also estimate models focusing on lead authors only stratified by sponsoring type. These models have the same model set-up and control variables as Model 6 in Table 4. The coefficients and confidence intervals of the regressions are shown in Fig. 3.

Regarding differences between sponsor types, Fig. 3 shows that the estimates of national competition are more negatively pronounced for industry-sponsored trials compared to non-industry sponsored trials, whereas estimates of local scientific reputation and local problem (DALY) are more positively pronounced for industry-sponsored trials compared to publicly-sponsored trials. It is noteworthy that the effect of local problem (DALY) is insignificant for trials not sponsored by industry, while the effect of scientific reputation is considerably less significant for trials not sponsored by industry. When we further explore these effects in regression models that include interaction terms for all main variables with industry sponsoring (not shown) we find a negative coefficient for the interaction term of industry sponsoring and national competition (β = -0.450, $p < 0.001$) and positive coefficients for the interaction terms of industry sponsoring with scientific reputation ($\beta = 1.716$, $p = 0.055$) and local problem ($\beta = 0.124$, $p = 0.010$).

Regarding differences between attribution of authorship to lead versus any author, we observe that the negative effect of distance to sponsor is somewhat more pronounced for lead authors compared to any authors, whereas the positive effect of scientific reputation is considerably more pronounced for lead authors compared to any author. When exploring differences between industry-sponsored and non-industry sponsored trials for the likelihood of becoming a lead author, we observe that estimates for local scientific reputation are considerable larger for industry-sponsored trials versus non-industry sponsored trials. This is confirmed by the estimated interaction term between industry-

Table 2
Correlation matrix.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]
[1] Authorship														
[2] Industry-sponsored	−0.25*													
[3] Distance to sponsor (log)	−0.19*	0.29*												
[4] National competition	−0.14*	−0.05*	−0.06*											
[5] Scientific reputation	0.23*	−0.10*	−0.13*	0.06*										
[6] Local problem (log)	−0.08*	0.09*	0.04*	0.11*	−0.06*									
[7] Total patients (log)	−0.26*	0.14*	0.10*	0.15*	−0.20*	0.30*								
[8] Duration (log)	0.03*	−0.34*	−0.08*	0.02*	−0.02*	−0.04*	0.30*							
[9] Phase 3	−0.16*	0.09*	0.11*	0.06*	−0.08*	0.08*	0.35*	0.14*						
[10] Small molecule	−0.02*	0.08*	0.03*	0.04*	−0.02*	0.12*	0.14*	−0.07*	0.01*					
[11] Total sites (log)	−0.39*	0.24*	0.16*	0.25*	−0.22*	0.26*	0.80*	0.28*	0.38*	0.07*				
[12] City sites (log)	0.08*	−0.04*	0.02*	−0.14*	0.12*	0.05*	0.14*	0.06*	0.03*	0.03*	0.18*			
[13] Trial experience (log)	0.21*	−0.06*	0.00	0.01*	0.65*	−0.05*	−0.23*	−0.06*	−0.08*	−0.02*	−0.22*	0.30*		
[14] Publication experience (log)	0.22*	−0.05*	−0.05*	−0.08*	0.86*	−0.07*	−0.18*	−0.01*	−0.07*	−0.02*	−0.21*	0.19*	0.68*	
[15] Sponsor headquarter	0.22*	−0.15*	−0.43*	−0.05*	0.12*	−0.03*	−0.09*	0.02*	−0.06*	0.00	−0.15*	0.06*	0.10*	0.10*

* $p < 0.01$.

Table 3
Logistic regressions of authorship at urban-area level.

	(1)	(2)	(3)
Industry-sponsor (H1)			−0.625***
			(0.039)
Trial controls			
Total patients		0.088***	0.079***
		(0.013)	(0.013)
Duration		0.263***	0.115***
		(0.022)	(0.020)
Phase 3		0.071***	0.065**
		(0.021)	(0.021)
Small molecule		−0.032	−0.014
		(0.023)	(0.023)
Total sites		−1.088***	−1.007***
		(0.020)	(0.020)
City controls			
City sites	0.043	0.773***	0.739***
	(0.044)	(0.038)	(0.039)
Trial experience	0.154***	0.053	0.051
	(0.033)	(0.032)	(0.032)
Publication experience	0.174***	0.134***	0.139***
	(0.020)	(0.017)	(0.017)
Sponsor headquarter	3.142***	2.557***	2.504***
	(0.270)	(0.244)	(0.240)
Constant	−3.746***	−0.816***	0.430*
	(0.137)	(0.199)	(0.210)
Disease dummies	No	Yes	Yes
Observations	112,503	112,503	112,503
Log likelihood	−45,944.3	−37,623.3	−37,372.6
LR test χ^2	9871***	26,513***	27,014***

Note: logistic regression with standard errors clustered across cities.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

sponsoring and scientific reputation which is positive and significant ($\beta = 5.523$, $p < 0.001$). Moreover, regarding differences between the estimates for local problem on becoming a lead author we observe a positively effect for industry-sponsoring ($\beta = 0.221$, $p = 0.065$).

5.5. Robustness checks

We conduct four additional analyses to check the robustness of our findings. First, as the local problem variable is defined on the national level we estimate a regression model on trial-publication pairs originating from international trials only. Second, as we find a positive effect of phase 3 trials on authorship likelihood and know that these trials are more data-intensive and follow strictly pre-determined protocols, we also estimate a regression model on a sample of trial-publication pairs originating from phase 3 trials only.

Third, instead of examining authorship attribution to a city on any corresponding publication, we estimate authorship attribution for each publication separately. We expect this to influence authorship likelihood estimates as some clinical trials have multiple corresponding publications with authorship numbers potentially differing across cities. This increases the number of observations in our sample to 227,929.

Fourth, we make an attempt to estimate authorship likelihood at the level of facilities. More specifically, we extract facility names from Clinicaltrials.gov and manually match these names with facility names on resulting publications. In this specification of our dependent variable an observation thus takes on a value of one in case the trial facility in the city is present on one of the corresponding publications, and zero otherwise. This makes our estimations more precise as authorship can no longer be attributed to researchers that are located in the same city but are not from the same facility. However, while adding granularity to our analysis, it also sharply reduces the number of observations as facility names are only available for half of the trials ($n = 1843$ trials and $n = 38.227$ city observations). Moreover, the trials in this subsample are considerably less often industry-sponsored (43.8 % versus 71.0 %) and smaller in terms of the number of cities involved (median of 9 versus 15).

Table 5 provides the results of the robustness checks. Model 1 on international trials finds similar effects for the local problem variable as well as for other variables, while the results for model 2 on phase 3 trials are also similar to earlier findings. Model 3 shows that estimating authorship likelihood for each publication separately does not change the results. In Model 4 we observe that overall effects are somewhat smaller than in previous models but the effects of distance to sponsor, national competition and scientific reputation remain significant. The local problem variable is no longer significant in this model.

Table 4
Conditional logit models of authorship at city level.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance to sponsor (H2)		−0.066***				−0.066***
		(0.012)				(0.011)
National competition (H3)			−0.367***			−0.545***
			(0.072)			(0.067)
Scientific reputation (H4)				2.697***		3.036***
				(0.730)		(0.736)
Local problem (H5)					0.131***	0.144***
					(0.029)	(0.027)
City sites	0.475***	0.480***	0.426***	0.496***	0.479***	0.436***
	(0.026)	(0.026)	(0.027)	(0.024)	(0.026)	(0.026)
Trial experience	0.145***	0.146***	0.158***	0.137***	0.141***	0.153***
	(0.027)	(0.027)	(0.026)	(0.027)	(0.027)	(0.025)
Publication experience	0.092***	0.092***	0.089***	0.055***	0.093***	0.046**
	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.016)
Sponsor headquarter	0.470***	0.194*	0.473***	0.451***	0.469***	0.173*
	(0.061)	(0.079)	(0.062)	(0.059)	(0.061)	(0.079)
Trial fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112,503	112,503	112,503	112,503	112,503	112,503
Log likelihood	−52,623.5	−52,590.1	−52,588.4	−52,582.3	−52,608.6	−52,474.8
LR test	4680.1***	4746.9***	4750.5***	4762.5***	4710.0***	4977.5***

Note: conditional logit models with clinical trial fixed effects and standard errors clustered across cities.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

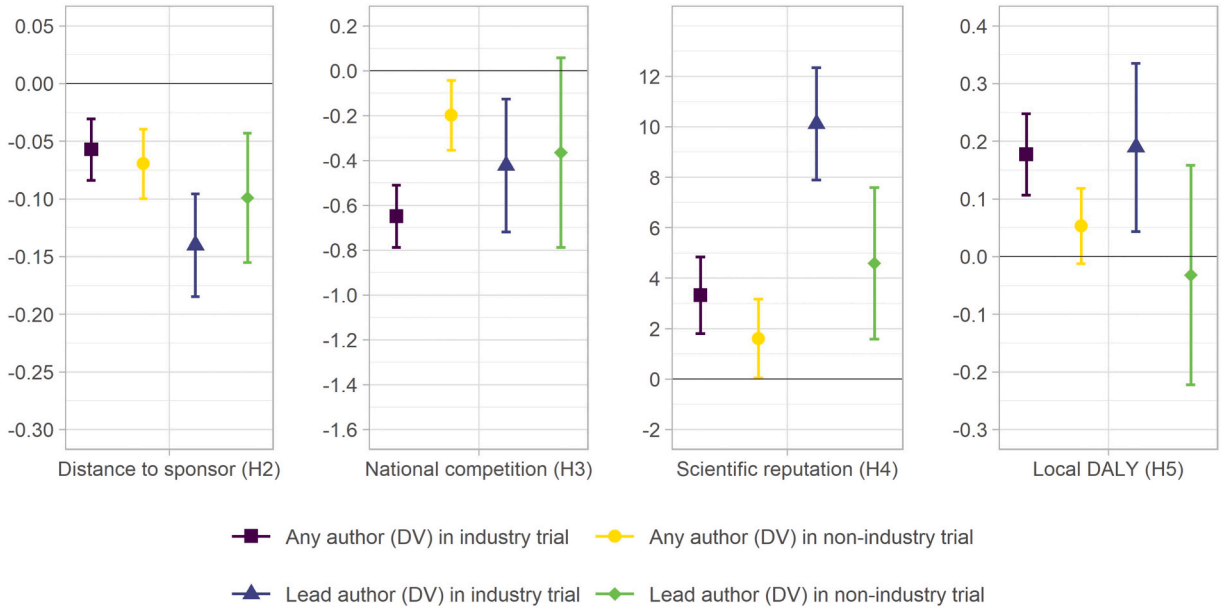


Fig. 3. Regression estimates per sponsor type and author type.
Note: all regressions have the same set-up as Model 6 in Table 4.

6. Discussion

The increasing prevalence of large geographically dispersed teams in scientific knowledge production has raised concerns over inequalities in authorship opportunities. Starting from a growing body of literature on the geography of scientific knowledge we developed and tested a number of hypotheses that consider how geography shapes such authorship opportunities for researchers involved in data collection tasks across the globe. Our analysis adds a geographic perspective to previous literature examining social determinants of authorship attribution (e.g., Haeussler and Sauermann, 2013; Jabbehdari and Walsh, 2017; Lissoni et al., 2020, 2013). Using a geographic lens, we confirm previous findings on the role of scientific experience and reputation as factors increasing opportunities for authorship, independent of contributions. We also add new understandings particularly with regard to how authorship opportunity structures are conditioned by geographic proximity as well as local reputation, demand and competition.

Our study adds to the literature on the geography of scientific knowledge which examines how locations and geographic proximity between researchers affect scientific knowledge production and diffusion. The positive effect of geographic proximity on establishing and maintaining research collaboration is well established (Catalini et al.,

Table 5
Robustness checks of conditional logit models.

	(1) International	(2) Phase 3	(3) All publications	(4) Facility level
Distance to sponsor (H2)	−0.052***	−0.065***	−0.069***	−0.035**
	(0.013)	(0.015)	(0.013)	(0.013)
National competition (H3)	−0.631***	−0.600***	−0.455***	−0.183*
	(0.061)	(0.081)	(0.080)	(0.077)
Scientific reputation (H4)	2.941***	3.166***	4.384***	2.728***
	(0.735)	(0.904)	(0.902)	(0.679)
Local problem (H5)	0.153***	0.180***	0.111***	0.032
	(0.028)	(0.037)	(0.032)	(0.035)
City sites	0.474***	0.407***	0.501***	0.477***
	(0.027)	(0.032)	(0.032)	(0.039)
Trial experience	0.158***	0.200***	0.166***	0.095***
	(0.028)	(0.029)	(0.027)	(0.027)
Publication experience	0.070***	0.055**	0.039*	0.013
	(0.018)	(0.018)	(0.019)	(0.019)
Sponsor headquarter	0.574***	0.420***	0.279**	0.012
	(0.092)	(0.120)	(0.086)	(0.064)
Trial fixed effects	Yes	Yes	Yes	Yes
Observations	94,372	76,139	227,929	38,227
Log likelihood	−40,407.8	−31,908.5	−72,239.8	−19,619.0
LR test	4367.8***	3656.5***	8616.8***	1080.0***

Note: Conditional logit models with clinical trial fixed effects and standard errors clustered across the city-level. Model 1 estimated on subset of only international trials; Model 2 estimated on subset of only phase 3 trials; Model 3 estimated for each trial-publication pair separately; Model 4 estimated on subset of trials with definition of authorship on facility level.

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

2020; Hoekman et al., 2009; Katz, 1994). Our results suggest that geographic proximity and underlying mechanisms relating to tacit knowledge exchange also matter for shaping authorship opportunities. Noteworthy are also the observed strong correlates of local research and publication experience with authorship attribution processes. Such observations suggest that local cumulative dynamics shape authorship opportunities in line with credit cycle dynamics and what has previously been described as a Matthew effect for locations (see e.g., Bonitz, 2005). Prior literature in the geography of science has also discussed the extent to which scientific knowledge production is directed to local problems and needs (Ciarli and Ràfols, 2019; Confraria and Wang, 2020). Our results show that authorship opportunities play a role in aligning research conduct with local needs.

We do find that authorship opportunity structures differ considerably depending on research project characteristics. Noteworthy is that while authorship networks of industry-sponsored studies are generally more closed, they also show more systematic geographic variation in authorship attribution, particularly with regard to the importance of local scientific reputation and demand and for becoming a lead author. These findings add to the literature on publication and authorship patterns of companies by showing spatial reflections of the logics and incentives that govern authorship attribution processes in industrial versus academic science (Sauermann and Stephan, 2013; Sismondo, 2009). In contrast, we observe that authorship attribution in non-industry-sponsored trials seem to be less related to local problems, reputation and competition dynamics. This may suggest that considerations of knowledge exchange and research capacity building through collaborative research have a higher priority in these projects.

Against this background, our findings have implications for current debates on authorship attribution in projects run by large, geographically dispersed research teams. Such implications relate to journal authorship policies as well as broader norms and practices in global team science to foster equitable authorship. Regarding the former, one implication relates to creating more explicit guidance on how authorship in large, geographically dispersed research projects should be established given known divisions-of-labor and specialized contributions. This includes establishing common and transparent standards for authorship across journals particularly with regard to how contributions to technical and data collection tasks translate into authorship and what is commonly regarded as substantial and major intellectual contributions in this regard. Our results suggest that implicit assumptions in authorship that rely on the primacy of conceptual contributions and manuscript drafting may contribute to exclusion and unequal translation of contributions into authorship across geographies, for instance shaped by geographic proximity to sponsors. Another measure would be to further standardise contributorship statements with explicit attention to technical tasks (cf. McNutt et al., 2018; Sauermann and Haeussler, 2017). This may result in more inclusive acknowledgement of all contributions and increase authorship opportunities for researchers who solely contributed to these tasks but would have difficulties contributing to others. The implementation of these proposals will require that academic journals stop arbitrarily limiting the number of authors of a manuscript or applying other policies that have similar effects (Morton et al., 2022).

Providing more clarity through revised authorship guidelines should be combined with discussions about social norms and practices in research teams with regard to how authorship is discussed and determined. Such discussions can be encouraged or even requested through journal policies. One approach would be to use more transparent authorship assignment systems in large geographically dispersed research teams such as those proposed by Whellan et al. (2015). This may reduce spatial and other biases and make authorship decisions more transparent. Following Morton et al. (2022), research teams may also use structured reflexivity statements for international research partnerships to encourage inclusive and open discussion on equitable authorship and broader issues of research prioritization and capacity strengthening through e.g., training and contributions to local infrastructure. It follows from our paper that one important reflection would be on whether researchers who contribute to data collection tasks are provided sufficient and equal opportunities to participate in activities that are deemed necessary to qualify for authorship as well as whether the various geographies involved are sufficiently represented through authorship.

It is, however, important to note that these initiatives will not necessarily solve more structural spatial inequalities in the science system. There is no guarantee that increasing transparency and clarifying authorship criteria contributes to broader normative goals of science to contribute to local capacity building (Rees et al., 2019). Inequalities might also be exacerbated when authorship attributions are strategically informed, for instance due to marketing reasons. Similarly, asking for more documentation on individual contributions in collaborative research may have unintended consequences. Researchers may shy away from collaborations that require additional bureaucracy and documentation. Open discussions around authorship attribution may be difficult if not impossible to have in very large research teams that are spread across different geographies (Smith et al., 2020). Finally, increasing transparency and clarity may not be sufficient to avoid all types of authorship conflicts that may arise through differences among the experiences, prestige, or power of contributors. These issues are not exclusively linked to collaborations across different geographies but may also occur in collaborations within a country, in interdisciplinary work, or in research characterized by a pronounced division of scientific labor.

There are several limitations to our study. One limitation is that we only observe whether local researchers contribute to data collection but do not know whether they also contribute to other tasks that are considered a requirement to qualify for authorship according to authorship guidelines. As discussed, such distinctions between technical and conceptual contributions are sometimes hard to make and often endogenous to the research process or alternatively, they might be hierarchical with researchers not knowing why they do or do not qualify for authorship. In practice, motivations and reasons for researchers to strive for authorship can also differ. Given these observations, we do not interpret our findings as revealing a misallocation of credit. Rather, the findings reveal systematic spatial variation in authorship opportunities in situations where such opportunities could have been equally provided to all due to involvement in data collection tasks. In the literature, these issues have been particularly discussed in relation to research activities conducted in LMICs. Our analysis was conducted on a global scale and future research may therefore study authorship attribution specifically for international collaborations between researchers from HICs and LMICs countries. These future studies may pay particular attention to further specifying what equitable authorship means in these contexts and how broader goals of capacity building and research prioritization in LMIC contexts can be realised.

A second limitation of our study is that while multi-city clinical trials are an illustrative case of large geographically dispersed research projects, the institutional set-up of clinical trials is specific. This makes it difficult to assess whether we can generalize our results to other fields. We do, however, increasingly observe large data-intensive research projects in other disciplines. As these projects share important characteristics with clinical trials such as an emphasis on detailed pre-defined

research protocols and geographically dispersed data collection efforts, we believe that our results are partially generalizable to other research projects and fields. Examples of settings include large data-intensive experimental studies in e.g., agronomy, development economics, educational sciences, and ecology as well as observational studies (e.g., sample collection, survey-based designs) in geography, earth sciences, oceanography, psychology as well as social sciences and economics in general. It is likely that the spatial mechanisms observed in this study also matter in other fields although they may play out differently depending on the specific context.

A third limitation of our study is that we cannot directly observe the rationale of why a specific trial site is selected to contribute to a clinical trial and whether authorship considerations played a role in the selection process. Future research could extend existing studies (Dombernowsky et al., 2019; Gehring et al., 2013) that have explored site-, investigator-, and locations-specific factors that influence the selection of clinical trial sites. An important question in our context is how considerations with regard to publications and authorship play a role in early phases of the research process including site selection and research design. Such questions can be related to broader ethical considerations that should be taken into account in international clinical trial selection and conduct (Glickman et al., 2009; Miller and Millum, 2022). Future studies could also pay attention to how choices regarding site selection and authorship attribution to local researchers impact the quality and legitimacy of research findings.

In all, our paper demonstrates that authorship opportunities in large geographically dispersed teams systematically vary depending on how researchers are spatially embedded in projects and local environments. Our findings signal unequal authorship opportunities for researchers across the globe which contributes to maintaining existing spatial inequalities in the science system. We endorse further research on the geography of authorship and initiatives by journal editors and the academic community to foster equitable authorship opportunities in large research teams independent of geographic location and as a fundamental principle for the conduct of science.

CRediT authorship contribution statement

Conceptualisation: J.H. (50 %) and B.R. (50 %); Methodology: J.H. (75 %) and B.R. (25 %); Data curation: J.H. (75 %) B.R. (25 %); Formal analysis: J.H. (100 %); Writing: J.H. (75 %) B.R. (25 %).

Declaration of competing interest

The authors declare that they have no competing interests.

Data availability

Data will be made available on request.

Acknowledgement

The authors would like to gratefully acknowledge comments received at the 10th Annual Workshop on Medical Innovation and Healthcare (WOMI), the 10th anniversary conference of the European Forum for Studies of Policies for Research and Innovation and the 18th International Joseph A. Schumpeter Society Conference 2021 as well as comments received from Koen Frenken on a draft version and useful advice from Jaap van Slageren.

Funding

The authors acknowledge that they received no external funding in support of this research.

Appendix 1. Urban area definition

The spatial dataset used for this study contains 1875 urban areas of the largest cities in the world. It uses the EU-OECD functional urban area (FUA) definition for urban areas in 38 EU and/or OECD countries and urban area definitions based on urban density and settlement points for other countries.

EU-OECD FUA

The EU-OECD functional urban area definition is used to define urban areas in a consistent way across EU and OECD countries. FUAs are composed of a densely inhabited city ('core') and a less densely populated commuting zone ('hinterland') whose labour market is highly integrated with the city (Dijkstra et al., 2019). The boundaries of FUAs are defined based on the presence of an urban area that consists of contiguous high-density population grids and municipality boundaries with at least half their population inside the urban area as well as a commuting zone based on commuting patterns (Dijkstra et al., 2019).

For this study we used the list of FUAs available in the January 2019 of the EU-OECD database covering 1191 FUAs in 34 countries (OECD 2019) plus 61 FUAs in 4 additional EU-countries (Eurostat 2018). FUAs are distributed over four groups: 1) small FUAs, population between 50,000 and 100,000, 2) medium-sized FUAs, population between 100,000 and 250,000, 3) Metropolitan FUAs population between 250,000 and 1.5 million, 4) Large metropolitan FUAs, with population above 1.5 million.

Other urban areas

To add urban areas in other countries, we rely on a general definition of an urban centre as there is no single consistent definition of FUAs across the world. We compiled a list of all cities with >500,000 inhabitants based on data derived from GeoNames Gazetteer[18] and the website City Population.[19] GeoNames Gazetter data is an open gazetteer database that includes geographic data such as place names, lat/long coordinates and population data retrieved from various sources. City Population is another open geographic database with population statistics for countries, administrative units, cities, urban areas and agglomerations mainly based on census data from national statistical. We included cities in case they had >500,000 inhabitants according to at least one of the two sources.

In order to create a spatial dataset of urban areas for these cities we relied on the Urban Extent Polygons from the Global Rural-Urban Mapping

[18] https://www.geonames.org/data-sources.html.
[19] http://citypopulation.de/references.html.

Project (CIESIN 2017). The database defines the spatial extent of urban areas based on population counts, night-time lights and (buffered) settlement points. Cities with >500,000 inhabitants were allocated to 571 urban areas and combined in case multiple cities were part of the same urban area. In a small number of cases the urban areas of cities were not clearly discernible from the Urban Extent Polygons or covered very large areas or multiple countries. In those cases, urban areas were defined based on the World Urban Areas, Landscan database of the Natural Earth Collection (Patterson, 2012). This led to further inclusion of 52 urban areas.

Results

Table A1.1 provides an overview of the results of the geocoding process and allocation of locations to urban areas. Out of 1,985,958 facilities in the entire ClinicalTrials.gov database, 86.5 % of participating facilities are allocated to one out of the 1.875 urban areas in the dataset. Fig. A1.1 provides an overview of the spatial distribution of clinical trials over urban areas.

Table A1.1
Allocation of facilities to urban areas.

	n (%)
Allocated to EU-OECD FUA	1,533,812 (77.2)
Allocated to other urban areas	183,955 (9.3)
Outside EU-OECD FUA or urban area	240,143 (12.1)
Not geocoded	28,048 (1.4)

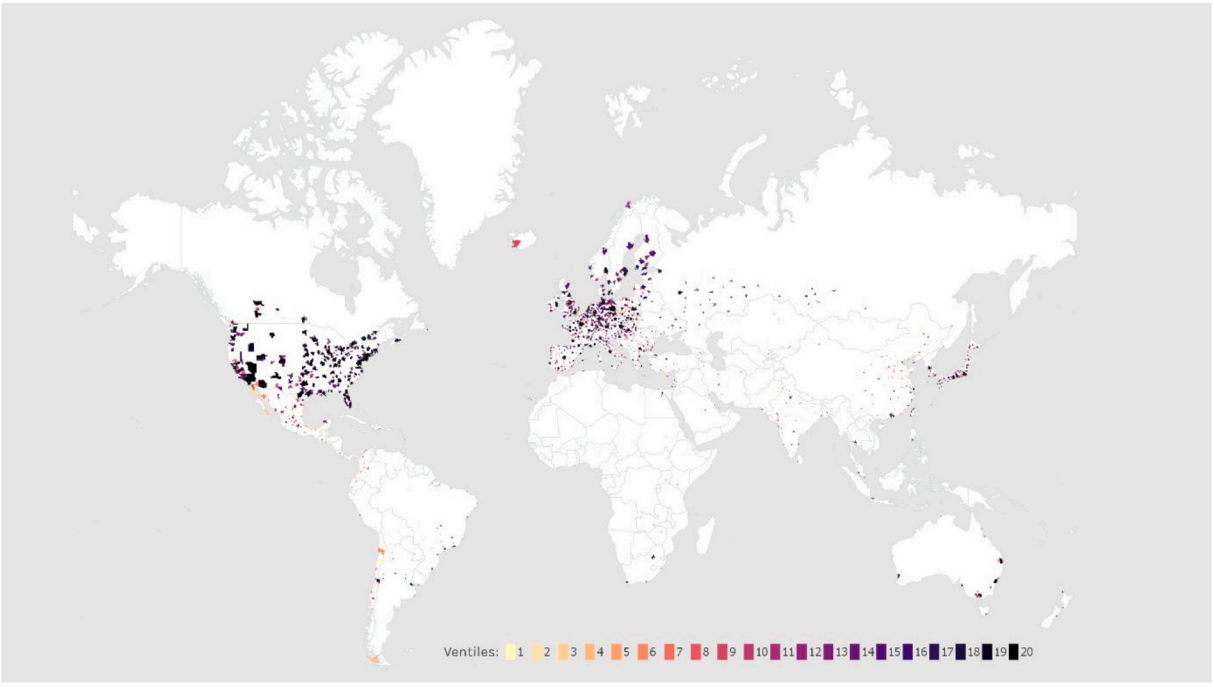
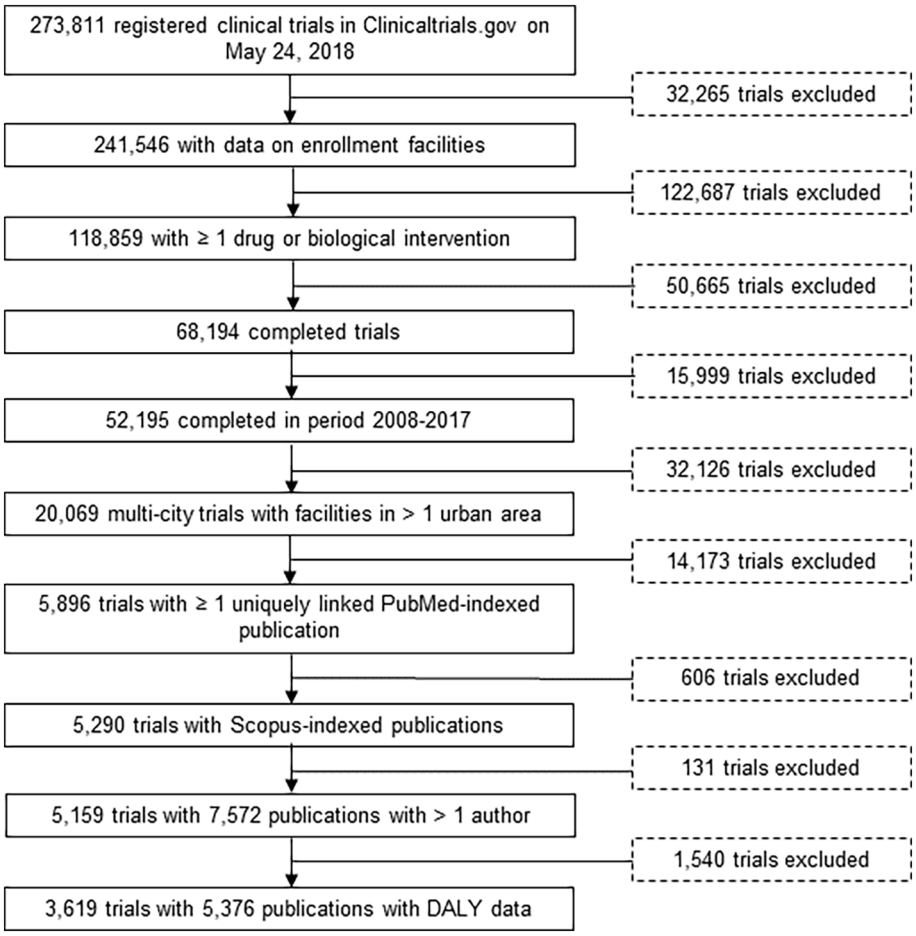


Fig. A1.1. Number of registered trials per urban area.

Appendix 2. Sample construction



Appendix 3. Disease categories

Table A3.1
Global health estimates level 2 cause categories.

Code	Cause name
20	Infectious and parasitic diseases
380	Respiratory infections
420	Maternal conditions
490	Neonatal conditions
540	Nutritional deficiencies
610	Malignant neoplasms
790	Other neoplasms
800	Diabetes mellitus
810	Endocrine, blood, immune disorders
820	Mental and substance use disorders
940	Neurological conditions
1020	Sense organ diseases
1100	Cardiovascular diseases
1170	Respiratory diseases
1210	Digestive diseases
1260	Genitourinary diseases
1330	Skin diseases
1340	Musculoskeletal diseases
1400	Congenital anomalies
1470	Oral conditions
1505	Sudden infant death syndrome
1520	Unintentional injuries
1600	Intentional injuries

Source: WHO (2018).

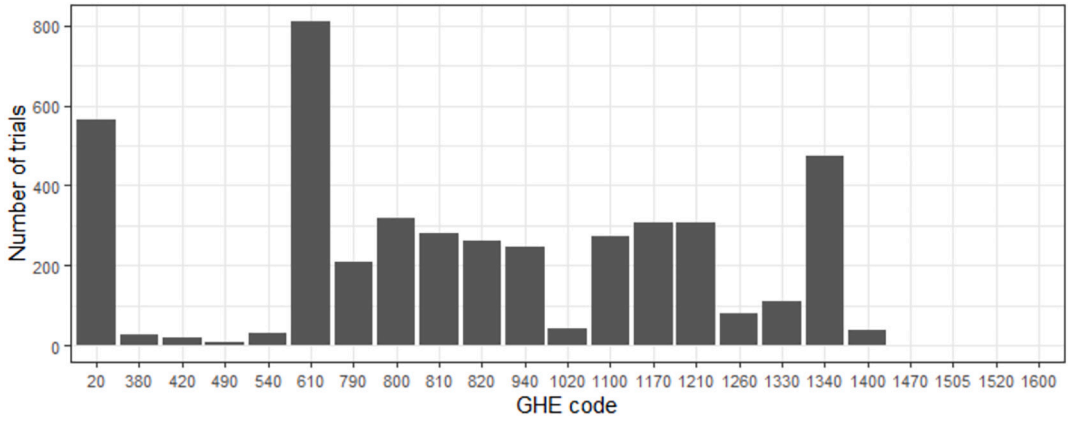


Fig. A3.1. Number of trials per disease area.
Note: GHE codes correspond with codes listed above. A trial can fit in multiple disease areas.

Appendix 4. Authorship rates per city and country

	Trial (n)	Author (n)	Rate (%)
Algeria	**10**	**1**	**0.10**
Alger	6	1	0.17
Oran	4	0	0.00
Argentina	**1061**	**97**	**0.09**
Buenos Aires	354	69	0.19
Cordoba	143	7	0.05
Guaymallen	87	1	0.01
Mardelplata	76	0	0.00
Rosario	152	8	0.05
Salta	35	0	0.00
San Miguel de Tucuman	115	9	0.08
Sanjuan	27	2	0.07
Santa Fe	72	1	0.01
Armenia	**2**	**1**	**0.50**
Yerevan	2	1	0.50
Australia	**2059**	**348**	**0.17**
Australian Capital Territory	49	1	0.02
Ballarat	10	0	0.00
Bendigo	5	0	0.00
Cairns	30	1	0.03
Geelong	58	4	0.07
Gold Coast	42	0	0.00
Greater Adelaide	279	42	0.15
Greater Brisbane	262	34	0.13
Greater Darwin	1	0	0.00
Greater Hobart	43	1	0.02
Greater Melbourne	427	126	0.30
Greater Perth	262	35	0.13
Greater Sydney	390	90	0.23
Newcastle	107	10	0.09
Sunshine Coast	53	3	0.06
Toowoomba	4	1	0.25
Townsville	1	0	0.00
Wollongong	36	0	0.00
Austria	**755**	**148**	**0.20**
Graz	124	12	0.10
Innsbruck	111	19	0.17
Klagenfurt	15	2	0.13
Linz	95	8	0.08
Salzburg	95	9	0.09
Vienna	315	98	0.31
Bangladesh	**5**	**4**	**0.80**
Chittagong	1	1	1.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Dhaka	4	3	0.75
Belarus	**35**	**1**	**0.03**
Gomel	10	0	0.00
Minsk	25	1	0.04
Belgium	**1705**	**323**	**0.19**
Antwerp	218	31	0.14
Bruges	87	9	0.10
Brussels	429	116	0.27
Charleroi	70	8	0.11
Gent	228	31	0.14
Kortrijk	53	2	0.04
Leuven	299	100	0.33
Liege	183	15	0.08
Mons	29	1	0.03
Namur	96	10	0.10
Ostend	13	0	0.00
Benin	**3**	**2**	**0.67**
Cotonou	3	2	0.67
Bosnia and Herzegovina	**16**	**2**	**0.13**
Sarajevo	16	2	0.13
Brazil	**1405**	**193**	**0.14**
Aracaju	4	0	0.00
Belem	20	1	0.05
Belo Horizonte	81	7	0.09
Brasilia	35	3	0.09
Campinas	91	4	0.04
Campo Grande	4	0	0.00
Coxipo Da Ponte	4	0	0.00
Curitiba	126	12	0.10
Fortaleza	43	2	0.05
Goiania	94	4	0.04
Itaquari	11	1	0.09
Joao Pessoa	3	0	0.00
Joinville	7	0	0.00
Juiz De Fora	20	0	0.00
Londrina	11	0	0.00
Maceio	5	0	0.00
Manaus	3	2	0.67
Natal	5	0	0.00
Porto Alegre	202	33	0.16
Recife	38	1	0.03
Ribeiraopreto	39	1	0.03
Riodejaneiro	171	36	0.21
Salvador	73	5	0.07
Sao Paolo	289	81	0.28
Sorocaba	15	0	0.00
Uberlandia	11	0	0.00
Bulgaria	**783**	**28**	**0.04**
Blagoevgrad	17	0	0.00
Burgas	19	0	0.00
Haskovo	6	0	0.00
Pazardzhik	18	0	0.00
Pleven	92	2	0.02
Plovdiv	119	5	0.04
Ruse	82	3	0.04
Shumen	8	0	0.00
Sliven	10	0	0.00
Sofia	218	14	0.06
Stara Zagora	42	1	0.02
Varna	106	3	0.03
Veliko Tarnovo	26	0	0.00
Vidin	9	0	0.00
Vratsa	7	0	0.00
Yambol	4	0	0.00
Burkina Faso	**8**	**7**	**0.88**
Bobo-Dioulasso	1	1	1.00
Ouagadougou	7	6	0.86
Cambodia	**1**	**1**	**1.00**
Phnum Penh	1	1	1.00
Cameroon	**1**	**1**	**1.00**
Yaounde	1	1	1.00
Canada	**4678**	**884**	**0.19**
Abbotsford	11	1	0.09
Brantford	3	0	0.00
Calgary	275	52	0.19
Edmonton	277	63	0.23

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Guelph	9	0	0.00
Halifax	210	30	0.14
Hamilton	267	60	0.22
Kitchener	112	17	0.15
London	219	29	0.13
Montreal	599	137	0.23
Niagara Falls	16	0	0.00
Ottawa	285	49	0.17
Peterborough	14	4	0.29
Quebec	311	35	0.11
Red Deer	12	0	0.00
Regina	40	4	0.10
Saanich	120	5	0.04
Saskatoon	117	6	0.05
Sherbrooke	119	5	0.04
St Catharines	23	0	0.00
St Johns	177	5	0.03
Toronto	680	256	0.38
Trois Rivieres	70	3	0.04
Vancouver	420	96	0.23
Windsor	70	4	0.06
Winnipeg	222	23	0.10
Chile	**386**	**25**	**0.06**
Antofagasta	3	0	0.00
Arica	1	0	0.00
Calama	1	0	0.00
Concepcion	24	0	0.00
Coquimbo-La Serena	3	0	0.00
Curico	2	0	0.00
Iquique	2	0	0.00
Los Angeles	1	0	0.00
Osorno	13	1	0.08
Puerto Montt	2	0	0.00
Punta Arenas	3	0	0.00
Quillota	10	0	0.00
Rancagua	17	0	0.00
Santiago	163	21	0.13
Talca	10	0	0.00
Temuco	34	1	0.03
Valdivia	24	1	0.04
Valparaiso	73	1	0.01
China	**2197**	**490**	**0.22**
Anshan	2	0	0.00
Bangbu	4	0	0.00
Baoding	2	0	0.00
Baotou	9	0	0.00
Beijing	156	85	0.54
Cangzhou	1	1	1.00
Changchun	30	6	0.20
Changsha	51	10	0.20
Changzhou	4	1	0.25
Chengdu	55	8	0.15
Chifeng	1	0	0.00
Chongqing	46	17	0.37
Dalian	18	2	0.11
Daqing	4	0	0.00
Fuzhou	21	6	0.29
Guilin	2	0	0.00
Guiyang	6	2	0.33
Haikou	13	0	0.00
Hangzhou	79	12	0.15
Harbin	27	4	0.15
Hefei	16	3	0.19
Hong Kong S.A.R.	147	49	0.33
Huaiyin	2	0	0.00
Huhehaote	2	0	0.00
Huizhou	1	1	1.00
Jiangyin	3	0	0.00
Jinan	36	6	0.17
Jingzhou	2	0	0.00
Jinzhou	2	0	0.00
Kaohsiung	140	16	0.11
Kunming	11	0	0.00
Lanzhou	7	1	0.14
Lianyungang	2	0	0.00
Mudanjiang	1	0	0.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Nanchang	17	1	0.06
Nanjing	66	16	0.24
Nanning	24	1	0.04
Nantong	1	0	0.00
Ningbo	2	0	0.00
Qingdao	21	2	0.10
Shanghai	145	66	0.46
Shantou	3	0	0.00
Shaoguan	1	1	1.00
Shenyang	49	4	0.08
Shenzhen	103	3	0.03
Shijiazhuang	29	3	0.10
Shiyan	2	0	0.00
Siping	3	2	0.67
Suzhou	34	8	0.24
Taian	1	0	0.00
Taichung	163	20	0.12
Tainan	90	20	0.22
Taipei	252	60	0.24
Taiyuan	15	1	0.07
Taizhou	1	0	0.00
Tianjin	72	20	0.28
Wenzhou	10	1	0.10
Wuhan	68	11	0.16
Wuhu	1	0	0.00
Wulumuqi	3	3	1.00
Wuxi	8	1	0.13
Xiamen	5	0	0.00
Xian	62	7	0.11
Xining	1	0	0.00
Xuzhou	2	0	0.00
Yancheng	1	0	0.00
Yangzhou	5	1	0.20
Yinchuan	9	2	0.22
Yueyang	2	0	0.00
Zhangjiakou	1	0	0.00
Zhanjiang	5	1	0.20
Zhengzhou	13	4	0.31
Zhenjiang	4	1	0.25
Colombia	**369**	**10**	**0.03**
Armenia	8	0	0.00
Barranquilla	73	0	0.00
Bogota D.C.	125	7	0.06
Bucaramanga	47	1	0.02
Cali	26	2	0.08
Cartagena	3	0	0.00
Ibague	1	0	0.00
Manizales	5	0	0.00
Medellin	64	0	0.00
Monteria	7	0	0.00
Neiva	1	0	0.00
Pereira	6	0	0.00
Rionegro	1	0	0.00
Yopal	1	0	0.00
Zipaquira	1	0	0.00
Costa Rica	**21**	**1**	**0.05**
San Jose	21	1	0.05
Croatia	**212**	**14**	**0.07**
Grad Zagreb	102	10	0.10
Osijek	33	0	0.00
Rijeka	41	2	0.05
Slavonski Brod	12	0	0.00
Split	23	2	0.09
Zadar	1	0	0.00
Cyprus	**3**	**0**	**0.00**
Lefkosia	3	0	0.00
Czech Republic	**1286**	**124**	**0.10**
Brno	205	28	0.14
Carlsbad	17	0	0.00
Ceske Budejovice	29	0	0.00
Chomutov	11	1	0.09
Hradec Kralove	124	17	0.14
Jihlava	17	0	0.00
Liberec	40	1	0.03
Most	2	0	0.00
Olomouc	107	9	0.08

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Ostrava	124	5	0.04
Pardubice	63	2	0.03
Plzen	84	3	0.04
Prague	378	57	0.15
Usti nad Labem	41	1	0.02
Zlin	44	0	0.00
Dem. Republic of Congo	**4**	**4**	**1.00**
Kinshasa	4	4	1.00
Denmark	**589**	**155**	**0.26**
Aalborg	88	9	0.10
Aarhus	125	30	0.24
Copenhagen	265	96	0.36
Odense	111	20	0.18
Dominican Republic	**15**	**3**	**0.20**
Santiago de los Caballeros	2	0	0.00
Santo Domingo	13	3	0.23
Ecuador	**21**	**5**	**0.24**
Guayaquil	10	3	0.30
Quito	11	2	0.18
Egypt	**61**	**10**	**0.16**
Alexandria	22	5	0.23
Cairo	36	5	0.14
El-Mahalla El-Kubra	2	0	0.00
Tanta	1	0	0.00
El Salvador	**3**	**0**	**0.00**
San Salvador	3	0	0.00
Estonia	**199**	**9**	**0.05**
Tallinn	115	5	0.04
Tartu	84	4	0.05
Finland	**612**	**72**	**0.12**
Helsinki	182	33	0.18
Jyvaskyla	41	0	0.00
Kuopio	69	5	0.07
Lahti	33	1	0.03
Oulu	83	6	0.07
Tampere	105	19	0.18
Turku	99	8	0.08
France	**5562**	**1292**	**0.23**
Ajaccio	4	0	0.00
Albi	11	0	0.00
Amiens	86	14	0.16
Angers	103	10	0.10
Angouleme	6	0	0.00
Annecy	26	1	0.04
Annemasse	1	0	0.00
Arras	15	0	0.00
Avignon	42	6	0.14
Bayonne	33	4	0.12
Beauvais	10	2	0.20
Belfort	11	1	0.09
Besancon	96	19	0.20
Beziers	13	0	0.00
Bordeaux	281	66	0.23
Boulogne-sur-Mer	14	1	0.07
Bourges	9	0	0.00
Brest	78	8	0.10
Brive-la-Gaillarde	10	0	0.00
Caen	118	30	0.25
Cannes	25	0	0.00
Chalons-en-Champagne	4	0	0.00
Chalon-sur-Saone	8	0	0.00
Chambery	12	0	0.00
Charleville-Mezieres	2	0	0.00
Chartres	12	0	0.00
Chateauroux	5	0	0.00
Cherbourg	6	0	0.00
Clermont-Ferrand	93	24	0.26
Colmar	33	1	0.03
Compiegne	7	0	0.00
Creil	4	0	0.00
Dijon	121	23	0.19
Douai	9	0	0.00
Dunkerque	11	1	0.09
Evreux	13	1	0.08
Fort-de-France	5	0	0.00
Frejus	7	0	0.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Grenoble	135	21	0.16
Henin - Carvin	4	0	0.00
La Rochelle	36	1	0.03
Le Havre	14	1	0.07
Le Mans	51	7	0.14
Lens - Lievin	11	1	0.09
Lille	277	80	0.29
Limoges	82	20	0.24
Lorient	12	2	0.17
Lyon	342	100	0.29
Marseille	267	66	0.25
Martigues	4	0	0.00
Melun	6	0	0.00
Metz	31	2	0.06
Montbeliard	15	1	0.07
Montpellier	240	53	0.22
Mulhouse	18	1	0.06
Nancy	168	44	0.26
Nantes	246	60	0.24
Nice	191	43	0.23
Nimes	64	6	0.09
Niort	6	0	0.00
Orleans	55	4	0.07
Paris	660	346	0.52
Pau	16	0	0.00
Perpignan	41	5	0.12
Poitiers	92	13	0.14
Quimper	7	0	0.00
Reims	101	19	0.19
Rennes	120	21	0.18
Roanne	9	1	0.11
Rouen	104	31	0.30
Saint-Brieuc	25	3	0.12
Saint-Etienne	73	9	0.12
Saint-Nazaire	11	1	0.09
Saint-Quentin	12	0	0.00
Strasbourg	198	28	0.14
Tarbes	9	1	0.11
Toulon	42	2	0.05
Toulouse	251	64	0.25
Tours	105	20	0.19
Troyes	4	1	0.25
Valence	11	1	0.09
Valenciennes	38	0	0.00
Vannes	14	1	0.07
Gabon	**2**	**2**	**1.00**
Libreville	2	2	1.00
Georgia	**30**	**4**	**0.13**
Tbilisi	30	4	0.13
Germany	**8108**	**1274**	**0.16**
Aachen	64	14	0.22
Aschaffenburg	63	2	0.03
Augsburg	53	2	0.04
Bamberg	33	1	0.03
Bayreuth	29	1	0.03
Berlin	695	128	0.18
Bielefeld	46	6	0.13
Bocholt, Stadt	6	0	0.00
Bonn	146	25	0.17
Brandenburg an der Havel	9	0	0.00
Braunschweig-Salzgitter Wolfsburg	42	0	0.00
Bremen	57	8	0.14
Bremerhaven	6	0	0.00
Celle	13	1	0.08
Chemnitz	31	0	0.00
Cologne	222	47	0.21
Constance	13	2	0.15
Cottbus	29	2	0.07
Darmstadt	26	2	0.08
Dessau	11	1	0.09
Dresden	279	44	0.16
Duren, Stadt	16	1	0.06
Dusseldorf	207	42	0.20
Erfurt	46	2	0.04
Flensburg	17	0	0.00
Frankfurt	10	0	0.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Frankfurt am Main	429	85	0.20
Freiburg im Breisgau	172	32	0.19
Friedrichshafen	7	0	0.00
Fulda	28	1	0.04
Gera	13	0	0.00
Giessen	48	4	0.08
Gorlitz	13	0	0.00
Gottingen	93	18	0.19
Greifswald	46	4	0.09
Halle an der Saale	86	6	0.07
Hamburg	467	103	0.22
Hanover	280	60	0.21
Heidelberg	212	40	0.19
Heilbronn	26	1	0.04
Hildesheim	47	0	0.00
Ingolstadt	12	0	0.00
Iserlohn	21	1	0.05
Jena	83	9	0.11
Kaiserslautern	11	2	0.18
Karlsruhe	47	3	0.06
Kassel	64	5	0.08
Kempten (Allgau)	9	0	0.00
Kiel	143	36	0.25
Koblenz	43	2	0.05
Krefeld	19	0	0.00
Landshut	12	0	0.00
Leipzig	254	22	0.09
Lubeck	89	7	0.08
Luneburg	10	0	0.00
Magdeburg	147	9	0.06
Mainz	222	67	0.30
Mannheim-Ludwigshafen	168	24	0.14
Marburg	95	14	0.15
Monchengladbach	28	0	0.00
Muenster	177	39	0.22
Munich	383	83	0.22
Neubrandenburg	5	1	0.20
Neumunster	16	0	0.00
Nuremberg	187	23	0.12
Offenburg	24	1	0.04
Oldenburg (Oldenburg)	65	8	0.12
Osnabruck	55	2	0.04
Paderborn	17	0	0.00
Passau	14	1	0.07
Pforzheim	7	0	0.00
Plauen	13	0	0.00
Regensburg	81	10	0.12
Remscheid	5	0	0.00
Reutlingen	6	1	0.17
Rosenheim	20	0	0.00
Rostock	72	6	0.08
Ruhr	405	75	0.19
Saarbrucken	89	7	0.08
Schweinfurt	14	0	0.00
Schwerin	34	1	0.03
Siegen	31	2	0.06
Solingen	20	2	0.10
Stralsund	16	1	0.06
Stuttgart	123	8	0.07
Trier	37	5	0.14
Tubingen	124	29	0.23
Ulm	145	34	0.23
Villingen-Schwenningen	22	3	0.14
Weimar	12	1	0.08
Wetzlar	13	0	0.00
Wiesbaden	71	16	0.23
Wilhelmshaven	12	0	0.00
Wuppertal	37	3	0.08
Wurzburg	130	26	0.20
Zwickau	13	0	0.00
Ghana	**2**	**1**	**0.50**
Kumasi	2	1	0.50
Greece	**530**	**58**	**0.11**
Athens	198	41	0.21
Chania	2	0	0.00
Ioannina	25	3	0.12

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Irakleio	55	2	0.04
Katerini	1	0	0.00
Larisa	52	3	0.06
Patras	47	1	0.02
Serres	5	0	0.00
Thessaloniki	142	8	0.06
Trikala	2	0	0.00
Volos	1	0	0.00
Guatamala	**37**	**2**	**0.05**
Ciudad de Guatemala	37	2	0.05
Guinea	**1**	**0**	**0.00**
Conakry	1	0	0.00
Haiti	**5**	**3**	**0.60**
Port-au-Prince	5	3	0.60
Honduras	**2**	**1**	**0.50**
San Pedro Sula	1	0	0.00
Tegucigalpa	1	1	1.00
Hungary	**1582**	**88**	**0.06**
Bekescsaba	27	0	0.00
Budapest	404	58	0.14
Debrecen	196	7	0.04
Dunaujvaros	8	0	0.00
Eger	37	0	0.00
Gyor	93	3	0.03
Kaposvar	45	1	0.02
Kecskemet	48	0	0.00
Miskolc	84	1	0.01
Nyiregyhaza	77	2	0.03
Pecs	108	1	0.01
Sopron	26	0	0.00
Szeged	135	6	0.04
Szekesfehervar	46	2	0.04
Szolnok	43	1	0.02
Szombathely	61	0	0.00
Tatabanya	22	0	0.00
Veszprem	63	5	0.08
Zalaegerszeg	59	1	0.02
Iceland	**15**	**4**	**0.27**
Reykjavik	15	4	0.27
India	**1622**	**100**	**0.06**
Agra	1	0	0.00
Ahmadabad	83	2	0.02
Ajmer	1	0	0.00
Aligarh	5	0	0.00
Allahabad	2	0	0.00
Amritsar	2	0	0.00
Aurangabad	10	0	0.00
Bangalore	170	17	0.10
Belgaum	13	0	0.00
Bhopal	7	0	0.00
Bhubaneswar	5	0	0.00
Bikaner	8	0	0.00
Chandigarh	16	1	0.06
Chennai	95	12	0.13
Coimbatore	48	1	0.02
Dehradun	2	0	0.00
Delhi	123	9	0.07
Durgapur	1	0	0.00
Firozabad	1	0	0.00
Gulbarga	2	0	0.00
Guntur	4	0	0.00
Guwahati	3	0	0.00
Hublidharwad	1	0	0.00
Hyderabad	130	6	0.05
Indore	46	5	0.11
Jaipur	68	1	0.01
Jalandhar	3	0	0.00
Jodhpur	1	0	0.00
Kannur	11	0	0.00
Kanpur	7	0	0.00
Kochi	45	0	0.00
Kolkata	49	4	0.08
Kozhikode	10	0	0.00
Lucknow	41	4	0.10
Ludhiana	33	1	0.03
Madurai	20	4	0.20

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Mangalore	46	0	0.00
Moradabad	1	0	0.00
Mumbai	134	14	0.10
Mysore	25	0	0.00
Nagpur	53	0	0.00
Nashik	26	1	0.04
Nellore	1	0	0.00
Patna	5	0	0.00
Pondicherry	2	2	1.00
Pune	135	12	0.09
Rajkot	3	0	0.00
Sangli	2	0	0.00
Surat	5	1	0.20
Thrissur	3	0	0.00
Tiruchirappalli	4	0	0.00
Tirunelveli	1	1	1.00
Trivandrum	39	1	0.03
Vadodara	19	0	0.00
Varanasi	14	1	0.07
Vijayawada	17	0	0.00
Visakhapatnam	20	0	0.00
Indonesia	**79**	**4**	**0.05**
Bandung	8	0	0.00
Banjarmasin	1	0	0.00
Denpasar	5	0	0.00
Jakarta	22	4	0.18
Malang	5	0	0.00
Medan	3	0	0.00
Padang	3	0	0.00
Palembang	3	0	0.00
Pekanbaru	1	0	0.00
Pontianak	1	0	0.00
Semarang	6	0	0.00
Surabaya	9	0	0.00
Surakarta	3	0	0.00
Ujungpandang	2	0	0.00
Yogyakarta	7	0	0.00
Iran	**6**	**5**	**0.83**
Shiraz	2	2	1.00
Tehran	4	3	0.75
Ireland	**194**	**31**	**0.16**
Cork	34	2	0.06
Dublin	118	28	0.24
Galway	25	1	0.04
Limerick	12	0	0.00
Waterford	5	0	0.00
Israel	**565**	**91**	**0.16**
Haifa	164	16	0.10
Jerusalem	144	18	0.13
Tel Aviv-Yafo	257	57	0.22
Italy	**4083**	**874**	**0.21**
Acireale	2	0	0.00
Alexandria	14	4	0.29
Altamura	2	0	0.00
Ancona	62	6	0.10
Andria	1	0	0.00
Arezzo	20	1	0.05
Asti	3	1	0.33
Avellino	18	2	0.11
Bari	75	9	0.12
Barletta	1	1	1.00
Battipaglia	1	1	1.00
Bergamo	83	21	0.25
Bologna	163	46	0.28
Bolzano	8	4	0.50
Brescia	99	30	0.30
Cagliari	74	16	0.22
Campobasso	6	0	0.00
Carpi	10	2	0.20
Caserta	13	2	0.15
Catania	96	9	0.09
Catanzaro	42	5	0.12
Cerignola	1	1	1.00
Como	20	4	0.20
Cosenza	17	1	0.06
Cremona	23	6	0.26

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Ferrara	55	8	0.15
Florence	156	34	0.22
Foggia	31	2	0.06
Forli	39	5	0.13
Gallarate	8	1	0.13
Genoa	190	36	0.19
Grosseto	10	2	0.20
La Spezia	7	1	0.14
L'Aquila	14	2	0.14
Latina	11	1	0.09
Lecce	24	3	0.13
Lecco	20	5	0.25
Livorno	12	3	0.25
Massa	5	1	0.20
Matera	4	1	0.25
Messina	37	3	0.08
Milan	458	155	0.34
Modena	69	14	0.20
Naples	178	38	0.21
Novara	43	13	0.30
Padua	124	26	0.21
Palermo	101	12	0.12
Parma	54	10	0.19
Pavia	119	18	0.15
Perugia	87	16	0.18
Pesaro	14	3	0.21
Pescara	28	6	0.21
Piacenza	24	4	0.17
Pisa	139	27	0.19
Pordenone	26	3	0.12
Potenza	16	3	0.19
Prato	22	3	0.14
Ragusa	6	5	0.83
Ravenna	31	4	0.13
Reggio di Calabria	28	9	0.32
Reggio nell`Emilia	51	9	0.18
Rimini	31	4	0.13
Rome	357	83	0.23
Salerno	8	2	0.25
Sassari	49	11	0.22
Sassuolo	1	1	1.00
Savona	2	0	0.00
Siracusa	5	1	0.20
Taranto	8	3	0.38
Terni	26	3	0.12
Toast	21	3	0.14
Trapani	2	1	0.50
Trent	7	2	0.29
Treviso	17	2	0.12
Trieste	26	6	0.23
Turin	194	46	0.24
Udine	62	19	0.31
Varese	30	5	0.17
Venice	19	3	0.16
Verona	86	15	0.17
Vicenza	37	6	0.16
Jamaica	**5**	**4**	**0.80**
Kingston	5	4	0.80
Japan	**2200**	**487**	**0.22**
Akita	11	3	0.27
Aomori	12	1	0.08
Asahikawa	22	0	0.00
Fuji	1	0	0.00
Fujieda	5	0	0.00
Fukui	15	1	0.07
Fukuoka	145	37	0.26
Fukushima	23	1	0.04
Hachinohe	21	0	0.00
Hakodate	5	2	0.40
Hamamatsu	28	2	0.07
Higashiosaka	216	92	0.43
Himeji	15	0	0.00
Hiroshima	71	11	0.15
Hitachi	17	0	0.00
Isesaki	7	1	0.14
Kagoshima	48	6	0.13

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Kanazawa	31	3	0.10
Kitakyushu	48	7	0.15
Kochi	25	2	0.08
Kofu	1	0	0.00
Koriyama	13	0	0.00
Kumamoto	81	14	0.17
Kurashiki	65	6	0.09
Kurume	29	2	0.07
Kusatsu	5	0	0.00
Kushiro	51	1	0.02
Marugame	3	0	0.00
Matsumoto	26	1	0.04
Matsuyama	30	6	0.20
Mito	6	0	0.00
Miyazaki	18	0	0.00
Morioka	19	0	0.00
Nagano	30	1	0.03
Nagasaki	42	7	0.17
Naha	45	4	0.09
Niigata	38	6	0.16
Numazu	13	0	0.00
Obihiro	8	0	0.00
Oita	43	4	0.09
Omuta	1	0	0.00
Sapporo	97	18	0.19
Sendai	56	7	0.13
Shimonoseki	1	0	0.00
Shizuoka	52	8	0.15
Shunan	3	0	0.00
Takamatsu	38	2	0.05
Takasaki	33	2	0.06
Tokushima	21	3	0.14
Tokyo	265	161	0.61
Tomakomai	8	0	0.00
Toyama	25	1	0.04
Toyohashi	9	1	0.11
Toyota	153	55	0.36
Ube	17	2	0.12
Utsunomiya	26	1	0.04
Wakayama	20	3	0.15
Yamagata	13	1	0.08
Yokkaichi	24	1	0.04
Yonago	6	0	0.00
Jordan	**8**	**1**	**0.13**
Amman	5	1	0.20
Irbid	3	0	0.00
Kazakhstan	**3**	**0**	**0.00**
Almaty	2	0	0.00
Chimkent	1	0	0.00
Kenya	**7**	**6**	**0.86**
Nairobi	7	6	0.86
Kuwait	**4**	**0**	**0.00**
Kuwait city	4	0	0.00
Latvia	**168**	**9**	**0.05**
Daugavpils	42	0	0.00
Jelgava	8	0	0.00
Liepaja	19	0	0.00
Riga	99	9	0.09
Lebanon	**28**	**8**	**0.29**
Beirut	28	8	0.29
Libya	**1**	**0**	**0.00**
Tripoli	1	0	0.00
Lithuania	**287**	**13**	**0.05**
Alytus	18	0	0.00
Kaunas	84	5	0.06
Klaipeda	57	0	0.00
Panevezys	9	0	0.00
Siauliai	29	0	0.00
Vilnius	90	8	0.09
Luxembourg	**4**	**0**	**0.00**
Luxembourg	4	0	0.00
Macedonia	**17**	**1**	**0.06**
Skopje	17	1	0.06
Malawi	**18**	**14**	**0.78**
Blantyre	10	6	0.60
Lilongwe	8	8	1.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Malaysia	**190**	**18**	**0.09**
Ipoh	19	0	0.00
Johor Bahru	26	0	0.00
Kota Bharu	24	5	0.21
Kuala Lumpur	101	13	0.13
Kuching	20	0	0.00
Mali	**6**	**5**	**0.83**
Bamako	6	5	0.83
Mexico	**1489**	**79**	**0.05**
Acapulco de Juarez	10	1	0.10
Aguascalientes	64	1	0.02
Cajeme	6	0	0.00
Celaya	5	0	0.00
Centro	12	0	0.00
Chihuahua	63	1	0.02
Coatzacoalcos	3	0	0.00
Colima	2	0	0.00
Cuautla	6	0	0.00
Cuernavaca	48	2	0.04
Culiacan	29	0	0.00
Durango	53	1	0.02
Ensenada	3	0	0.00
Guadalajara	250	16	0.06
Guadalupe	1	0	0.00
Hermosillo	14	0	0.00
Irapuato	1	0	0.00
Juarez	4	0	0.00
Leon	33	1	0.03
Los Cabos	1	0	0.00
Merida	56	4	0.07
Mexicali	20	0	0.00
Mexico City	284	32	0.11
Monclova	1	0	0.00
Monterrey	225	12	0.05
Morelia	36	2	0.06
Oaxaca de Juarez	7	0	0.00
Pachuca de Soto	21	0	0.00
Puebla	24	1	0.04
Puerto Vallarta	1	0	0.00
Queretaro	20	0	0.00
Saltillo	9	0	0.00
San Juan del Rio	1	0	0.00
San Luis Potosi	85	2	0.02
Tampico	20	1	0.05
Tapachula	1	0	0.00
Tijuana	22	1	0.05
Toluca	20	1	0.05
Torreon	8	0	0.00
Veracruz	10	0	0.00
Xalapa	10	0	0.00
Moldova	**10**	**2**	**0.20**
Chisinau	10	2	0.20
Morocco	**12**	**0**	**0.00**
Casablanca	4	0	0.00
Marrakech	2	0	0.00
Meknes	1	0	0.00
Rabat	5	0	0.00
Mozambique	**2**	**0**	**0.00**
Maputo	2	0	0.00
Myanmar	**3**	**1**	**0.33**
Mandalay	1	0	0.00
Yangon	2	1	0.50
Netherlands	**1628**	**297**	**0.18**
Alkmaar	34	2	0.06
Almelo	30	0	0.00
Alphen aan den Rijn	3	0	0.00
Amersfoort	34	3	0.09
Amsterdam	241	80	0.33
Apeldoorn	18	1	0.06
Arnhem	42	6	0.14
Assen	9	1	0.11
Bergen op Zoom	3	0	0.00
Breda	62	4	0.06
Deventer	16	1	0.06
Ede	14	1	0.07
Eindhoven	91	4	0.04

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Enschede	48	1	0.02
Gouda	21	2	0.10
Greater Soest	1	0	0.00
Groningen	107	29	0.27
Heerlen	46	5	0.11
Leeuwarden	30	2	0.07
Leiden	70	22	0.31
Lelystad	2	0	0.00
Maastricht	58	13	0.22
Middelburg	2	0	0.00
Nijmegen	113	30	0.27
Oss	2	1	0.50
Roosendaal	5	0	0.00
Rotterdam	179	34	0.19
s-Hertogenbosch	34	5	0.15
Sittard-Geleen	39	1	0.03
The Hague	74	3	0.04
Tilburg	36	1	0.03
Utrecht	117	40	0.34
Venlo	13	2	0.15
Zwolle	34	3	0.09
New Zealand	**338**	**55**	**0.16**
Auckland	158	40	0.25
Christchurch	112	11	0.10
Wellington	68	4	0.06
Niger	**1**	**0**	**0.00**
Niamey	1	0	0.00
Nigeria	**3**	**2**	**0.67**
Ibadan	1	1	1.00
Ilorin	1	1	1.00
Jos	1	0	0.00
Norway	**355**	**79**	**0.22**
Bergen	67	14	0.21
Kristiansand	11	1	0.09
Oslo	136	39	0.29
Stavanger	55	12	0.22
Tromso	25	2	0.08
Trondheim	61	11	0.18
Oman	**1**	**0**	**0.00**
Muscat	1	0	0.00
Pakistan	**21**	**2**	**0.10**
Karachi	10	0	0.00
Lahore	6	1	0.17
Multan	2	0	0.00
Rawalpindi	3	1	0.33
Peru	**183**	**27**	**0.15**
Chiclayo	4	0	0.00
Paucarpata	27	0	0.00
San Juan de Lurigancho	147	27	0.18
Trujillo	5	0	0.00
Philippines	**270**	**29**	**0.11**
Angeles	3	0	0.00
Cagayan de Oro	1	0	0.00
Cebu	56	3	0.05
Davao	29	1	0.03
Iloilo	38	0	0.00
Manila	143	25	0.17
Poland	**3327**	**252**	**0.08**
Bialystok	226	9	0.04
Bielsko-Biala	25	0	0.00
Bydgoszcz	155	4	0.03
Chelm	1	0	0.00
Cracow	291	17	0.06
Czestochowa	22	0	0.00
Elblag	66	1	0.02
Elk	2	0	0.00
Gdansk	250	13	0.05
Gorzow Wielkopolski	11	0	0.00
Grudziadz	15	0	0.00
Inowroclaw	9	0	0.00
Jastrzebie Zdroj	3	0	0.00
Jelenia Gora	5	0	0.00
Kalisz	7	0	0.00
Katowice	237	19	0.08
Kielce	45	2	0.04
Konin	3	3	1.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Koszalin	7	0	0.00
Legnica	5	0	0.00
Leszno	4	0	0.00
Lodz	253	43	0.17
Lomza	6	0	0.00
Lubin	17	0	0.00
Lublin	232	12	0.05
Nowy Sacz	1	0	0.00
Olsztyn	49	0	0.00
Opole	17	0	0.00
Ostrow Wielkopolski	20	0	0.00
Ostrowiec Swietokrzyski	6	0	0.00
Pabianice	4	0	0.00
Pila	8	0	0.00
Piotrkow Trybunalski	7	0	0.00
Plock	17	0	0.00
Poznan	216	16	0.07
Przemysl	2	0	0.00
Radom	12	0	0.00
Rybnik	9	0	0.00
Rzeszow	35	2	0.06
Siedlce	5	0	0.00
Slupsk	10	1	0.10
Stalowa Wola	4	0	0.00
Suwalki	1	0	0.00
Swidnica	3	0	0.00
Szczecin	108	5	0.05
Tarnow	45	0	0.00
Tczew	6	0	0.00
Tomaszow Mazowiecki	1	0	0.00
Torun	85	0	0.00
Walbrzych	7	0	0.00
Warsaw	474	78	0.16
Wloclawek	10	0	0.00
Wroclaw	253	27	0.11
Zamosc	10	0	0.00
Zielona Gora	5	0	0.00
Portugal	**412**	**31**	**0.08**
Aveiro	14	1	0.07
Braga	11	0	0.00
Coimbra	85	8	0.09
Faro	16	0	0.00
Funchal	1	0	0.00
Guimaraes	4	0	0.00
Lisbon	161	17	0.11
Ponta Delgada	3	0	0.00
Porto	102	5	0.05
Viana do Castelo	4	0	0.00
Vila Franca de Xira	5	0	0.00
Viseu	6	0	0.00
Qatar	**3**	**1**	**0.33**
Doha	3	1	0.33
Romania	**1364**	**56**	**0.04**
Alba Iulia	8	0	0.00
Arad	18	0	0.00
Bacau	25	0	0.00
Baia Mare	33	2	0.06
Bistrita	1	0	0.00
Botosani	1	0	0.00
Brasov	94	3	0.03
Bucuresti	317	25	0.08
Buzau	7	0	0.00
Calarasi	1	1	1.00
Cluj-Napoca	134	8	0.06
Constanta	50	4	0.08
Craiova	59	2	0.03
Focsani	10	0	0.00
Galati	38	1	0.03
Iasi	141	5	0.04
Oradea	61	0	0.00
Piatra Neamt	2	0	0.00
Pitesti	24	0	0.00
Ploiesti	36	0	0.00
Ramnicu Valcea	4	0	0.00
Roman	1	0	0.00
Satu Mare	10	0	0.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Sibiu	52	1	0.02
Slatina	1	0	0.00
Suceava	19	0	0.00
Târgoviste	12	0	0.00
Târgu Mures	97	0	0.00
Timisoara	108	4	0.04
Russia	**2624**	**152**	**0.06**
Astrachan	2	0	0.00
Barnaul	76	1	0.01
Celjabinsk	65	2	0.03
Habarovsk	2	0	0.00
Irkutsk	24	0	0.00
Izevsk	20	1	0.05
Jaroslavl	157	0	0.00
Jekaterinburg	100	0	0.00
Kazan	136	1	0.01
Kemerovo	67	1	0.01
Kirov	25	2	0.08
Krasnodar	37	5	0.14
Krasnojarsk	31	0	0.00
Lipeck	16	0	0.00
Moscow	481	80	0.17
Niznij Novgorod	107	3	0.03
Novo Kuzneck	4	0	0.00
Novosibirsk	126	0	0.00
Omsk	23	2	0.09
Orenburg	20	0	0.00
Penza	31	1	0.03
Perm	31	0	0.00
Rjazan	57	2	0.04
Rostov	53	1	0.02
Samara	100	5	0.05
Sankt Peterburg	445	36	0.08
Saratov	116	2	0.02
Tjumen	33	0	0.00
Toljatti	1	1	1.00
Tomsk	78	1	0.01
Ufa	54	5	0.09
Uljanovsk	19	0	0.00
Vladivostok	3	0	0.00
Volgograd	36	0	0.00
Voronez	48	0	0.00
Rwanda	**4**	**4**	**1.00**
Kigali	4	4	1.00
Saudi Arabia	**31**	**5**	**0.16**
Ad Damman	5	0	0.00
Jeddah	6	0	0.00
Riyadh	20	5	0.25
Senegal	**3**	**3**	**1.00**
Dakar	3	3	1.00
Serbia	**101**	**10**	**0.10**
Beograd	101	10	0.10
Singapore	**126**	**32**	**0.25**
Singapore	126	32	0.25
Slovakia	**518**	**16**	**0.03**
Banska Bystrica	52	1	0.02
Bratislava	165	8	0.05
Kosice	103	2	0.02
Nitra	60	4	0.07
Presov	46	0	0.00
Trencin	18	1	0.06
Trnava	28	0	0.00
Zilina	46	0	0.00
Slovenia	**48**	**5**	**0.10**
Ljubljana	36	5	0.14
Maribor	12	0	0.00
South Africa	**772**	**115**	**0.15**
Cape Town	246	45	0.18
Durban	165	17	0.10
Johannesburg	296	52	0.18
Pietermaritzburg	8	0	0.00
Port Elizabeth	57	1	0.02
South Korea	**844**	**220**	**0.26**
Chuncheon	7	1	0.14
Dalseong	114	17	0.15
Deokjin	29	4	0.14

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Gangneung	3	2	0.67
Gimhae	106	18	0.17
Gwangsan	60	7	0.12
Heungdeok	22	1	0.05
Iksan	3	1	0.33
Jeju	3	0	0.00
Jinju	10	2	0.20
Naju	1	1	1.00
Nam	14	5	0.36
Sebuk	7	3	0.43
Seo	51	8	0.16
Seongsan	3	1	0.33
Seoul	378	147	0.39
Wonju	33	2	0.06
Spain	**3959**	**665**	**0.17**
Albacete	10	3	0.30
Alcoy	3	0	0.00
Alicante	116	9	0.08
Almeria	32	1	0.03
Avila	9	0	0.00
Aviles	5	0	0.00
Badajoz	21	0	0.00
Barcelona	633	226	0.36
Basin	4	0	0.00
Benidorm	10	0	0.00
Bilbao	104	5	0.05
Burgos	23	1	0.04
Caceres	31	5	0.16
Cadiz	26	3	0.12
Cartagena	9	1	0.11
Castellon de la Plana	14	2	0.14
Ceuta	3	0	0.00
Ciudad Real	13	3	0.23
Cordoba	78	10	0.13
Coruna (A)	114	14	0.12
Donostia-San Sebastian	48	9	0.19
Eivissa	2	0	0.00
Elche/Elx	32	0	0.00
Elda	9	0	0.00
Ferrol	9	1	0.11
Gandia	5	0	0.00
Gijon	13	0	0.00
Girona	55	5	0.09
Granada	61	5	0.08
Guadalajara	24	1	0.04
Huelva	9	1	0.11
Igualada	1	1	1.00
Jaen	23	1	0.04
Jerez de la Frontera	17	1	0.06
Las Palmas	21	1	0.05
Leon	18	2	0.11
Linea de la Concepcion, La	1	0	0.00
Lleida	35	5	0.14
Logrono	7	1	0.14
Lugo	21	1	0.05
Madrid	566	134	0.24
Malaga	158	18	0.11
Manresa	10	0	0.00
Marbella	12	0	0.00
Melilla	1	0	0.00
Merida	27	0	0.00
Murcia	42	4	0.10
Ourense	11	1	0.09
Oviedo	78	8	0.10
Palencia	3	0	0.00
Palma de Mallorca	100	12	0.12
Pamplona	68	9	0.13
Ponferrada	14	0	0.00
Pontevedra	22	2	0.09
Puerto de la Cruz	1	0	0.00
Reus	21	2	0.10
Sagunto	16	0	0.00
Salamanca	66	19	0.29
Sanlucar de Barrameda	8	0	0.00
Santa Cruz de Tenerife	47	6	0.13
Santander	104	11	0.11

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Santiago de Compostela	130	15	0.12
Saragossa	68	10	0.15
Seville	261	25	0.10
Tarragona	24	3	0.13
Toledo	15	0	0.00
Torrevieja	10	0	0.00
Valencia	294	57	0.19
Valladolid	42	4	0.10
Vigo	43	5	0.12
Vitoria	15	1	0.07
Zamora	13	1	0.08
Sri Lanka	**3**	**1**	**0.33**
Colombo	3	1	0.33
Sweden	**1002**	**201**	**0.20**
Boras	14	0	0.00
Gothenburg	179	42	0.23
Helsingborg	22	2	0.09
Jonkoping	21	1	0.05
Linkoping	70	6	0.09
Malmo	184	39	0.21
Norrkoping	8	0	0.00
Orebro	51	3	0.06
Stockholm	264	77	0.29
Umea	69	9	0.13
Uppsala	108	22	0.20
Vasteras	12	0	0.00
Switzerland	**586**	**154**	**0.26**
Basel	81	45	0.56
Bern	92	25	0.27
Biel/Bienne	11	0	0.00
Geneve	68	14	0.21
Lausanne	72	17	0.24
Lucerne	13	2	0.15
Lugano	44	3	0.07
St. Gallen	66	14	0.21
Winterthur	10	4	0.40
Zurich	129	30	0.23
Tanzania	**2**	**2**	**1.00**
Dar es Salaam	2	2	1.00
Thailand	**153**	**34**	**0.22**
Bangkok	153	34	0.22
Togo	**1**	**0**	**0.00**
Lome	1	0	0.00
Tunisia	**18**	**2**	**0.11**
Tunis	18	2	0.11
Turkey	**603**	**61**	**0.10**
Adana	39	4	0.10
Adapazari	1	0	0.00
Ankara	125	17	0.14
Antalya	36	2	0.06
Bursa	26	3	0.12
Denizli	9	0	0.00
Diyarbakir	8	1	0.13
Eskisehir	13	2	0.15
Gaziantep	20	0	0.00
Icel	22	1	0.05
Istanbul	150	18	0.12
Izmir	102	12	0.12
Kahramanmaras	5	0	0.00
Kayseri	18	0	0.00
Konya	14	1	0.07
Samsun	15	0	0.00
Uganda	**18**	**18**	**1.00**
Kampala	18	18	1.00
Ukraine	**1063**	**39**	**0.04**
Dnipropetrovsk	149	8	0.05
Donetsk	131	6	0.05
Kharkiv	187	4	0.02
Kryvyi Rih	11	0	0.00
Kyiv	240	12	0.05
Lviv	119	7	0.06
Mykolayiv	17	0	0.00
Odesa	109	0	0.00
Zaporizhzhya	100	2	0.02
United Arab Emirates	**8**	**0**	**0.00**
Abu Dhabi	4	0	0.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Al Ayn	1	0	0.00
Dubayy	3	0	0.00
United Kingdom	**3801**	**715**	**0.19**
Aberdeen	41	2	0.05
Ashford	9	1	0.11
Basingstoke and Deane	17	3	0.18
Bath and North East Somerset	49	1	0.02
Bedford	4	0	0.00
Belfast	76	12	0.16
Blackburn with Darwen	9	0	0.00
Blackpool	34	0	0.00
Bournemouth	41	3	0.07
Bracknell Forest	1	0	0.00
Brighton and Hove	51	9	0.18
Bristol	97	14	0.14
Burnley	1	0	0.00
Cambridge	116	21	0.18
Cannock Chase	16	0	0.00
Cardiff	74	13	0.18
Carlisle	2	0	0.00
Cheltenham	6	1	0.17
Cheshire West and Chester	14	0	0.00
Chesterfield	26	0	0.00
Colchester	12	1	0.08
Corby	13	0	0.00
Coventry	75	5	0.07
Crawley	11	0	0.00
Darlington	7	0	0.00
Derby	23	3	0.13
Derry & Strabane	7	0	0.00
Doncaster	8	0	0.00
Dundee City	55	7	0.13
East Staffordshire	4	1	0.25
Eastbourne	1	0	0.00
Edinburgh	96	15	0.16
Exeter	33	3	0.09
Falkirk	3	0	0.00
Glasgow	177	32	0.18
Gloucester	4	0	0.00
Great Yarmouth	4	1	0.25
Guildford	51	5	0.10
Hartlepool	1	0	0.00
Hastings	17	0	0.00
Ipswich	12	0	0.00
Kettering	2	0	0.00
Kingston upon Hull	45	4	0.09
Leeds	147	33	0.22
Leicester	89	19	0.21
Lincoln	7	0	0.00
Liverpool	138	22	0.16
London	556	269	0.48
Luton	4	0	0.00
Maidstone	21	1	0.05
Manchester	258	44	0.17
Mansfield	1	0	0.00
Medway	8	0	0.00
Middlesbrough	43	1	0.02
Milton Keynes	2	0	0.00
Newcastle upon Tyne	148	20	0.14
Newport	8	0	0.00
North East Lincolnshire	1	0	0.00
Northampton	21	0	0.00
Norwich	28	2	0.07
Nottingham	107	17	0.16
Nuneaton and Bedworth	11	0	0.00
Oxford	82	32	0.39
Peterborough	16	0	0.00
Plymouth	75	5	0.07
Portsmouth	18	1	0.06
Preston	11	1	0.09
Reading	23	0	0.00
Redditch	1	0	0.00
Rushmoor	7	1	0.14
Sheffield	112	20	0.18
Slough	4	0	0.00
Southampton	76	14	0.18

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Stevenage	15	1	0.07
Stoke-on-Trent	39	0	0.00
Sunderland	13	0	0.00
Swansea	37	8	0.22
Swindon	9	0	0.00
Telford and Wrekin	3	0	0.00
Thanet	4	0	0.00
Torbay	22	2	0.09
Warrington	2	0	0.00
West Midlands urban area	220	39	0.18
Wirral	26	3	0.12
Woking	2	0	0.00
Worcester	4	0	0.00
Worthing	4	0	0.00
Wrexham	6	1	0.17
Wycombe	4	0	0.00
York	23	2	0.09
United States	**42,826**	**8113**	**0.19**
Ada	179	3	0.02
Alachua	230	38	0.17
Albany	188	19	0.10
Albuquerque	244	20	0.08
Allen	46	1	0.02
Atlanta	841	176	0.21
Atlantic City	41	1	0.02
Austin	342	36	0.11
Bell	87	2	0.02
Benton (AR)	11	1	0.09
Benton (MN)	25	0	0.00
Benton (WA)	39	0	0.00
Berks	116	2	0.02
Boston	750	428	0.57
Boulder	68	0	0.00
Brazos	15	1	0.07
Brevard	88	1	0.01
Broome	58	0	0.00
Brown	42	0	0.00
Butte	5	0	0.00
Caddo	144	7	0.05
Cameron	14	1	0.07
Cass	102	3	0.03
Centre	29	0	0.00
Champaign	41	0	0.00
Charleston	394	57	0.14
Charlotte	370	20	0.05
Chatham	120	4	0.03
Chicago	985	319	0.32
Cincinnati	599	118	0.20
Collier	46	1	0.02
Columbus	445	91	0.20
Comanche	9	0	0.00
Cumberland (ME)	66	6	0.09
Cumberland (NC)	28	1	0.04
Cuyahoga	558	159	0.28
Dallas	845	191	0.23
Dane	189	50	0.26
Dauphin	137	13	0.09
Davidson	437	110	0.25
Delaware	30	0	0.00
Denver	694	171	0.25
Detroit (Greater)	586	98	0.17
Douglas (KS)	18	0	0.00
Douglas (NE)	348	39	0.11
Durham	529	249	0.47
East Baton Rouge	107	11	0.10
Ector	40	1	0.03
El Paso (CO)	165	4	0.02
El Paso (TX)	96	10	0.10
Erie (NY)	255	36	0.14
Erie (PA)	52	0	0.00
Escambia	75	1	0.01
Fayette	242	23	0.10
Flagler-Daytona Beach	1	0	0.00
Forsyth	317	52	0.16
Fresno (Greater)	160	9	0.06
Genesee	83	0	0.00

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Greene	92	9	0.10
Greenville	279	8	0.03
Guilford	171	4	0.02
Hamilton (TN)	138	9	0.07
Hampden	95	9	0.09
Harrison	28	0	0.00
Hartford	170	18	0.11
Hidalgo	40	0	0.00
Houston	973	286	0.29
Indian River	60	2	0.03
Indianapolis	523	113	0.22
Ingham	89	4	0.04
Jackson (MO)	474	68	0.14
Jackson (OR)	110	12	0.11
Jacksonville	455	63	0.14
Jefferson (AL)	573	124	0.22
Jefferson (KY)	271	20	0.07
Jefferson (TX)	22	1	0.05
Johnson	171	40	0.23
Kalamazoo	133	2	0.02
Kankakee	10	0	0.00
Kent	97	7	0.07
Kern	57	6	0.11
Knox	138	5	0.04
Lackawanna	26	1	0.04
Lafayette	74	0	0.00
Lafayette (IN)	25	1	0.04
Lancaster (NE)	131	6	0.05
Lancaster (PA)	52	2	0.04
Lane	104	5	0.05
Larimer	66	2	0.03
Las Cruces	12	0	0.00
Las Vegas	373	35	0.09
Lee	91	4	0.04
Lehigh	148	10	0.07
Linn	33	0	0.00
Los Angeles (Greater)	1363	426	0.31
Lubbock	79	2	0.03
Lucas	230	7	0.03
Luzerne	33	1	0.03
Madison	127	3	0.02
Mahoning	24	0	0.00
Marion (FL)	117	1	0.01
Marion (OR)	28	1	0.04
McLean	58	3	0.05
McLennan	70	7	0.10
Memphis	354	49	0.14
Merced	14	0	0.00
Mesa	15	0	0.00
Miami (Greater)	962	131	0.14
Midland	26	1	0.04
Milwaukee	328	41	0.13
Minneapolis	510	111	0.22
Minnehaha	66	2	0.03
Mobile	168	3	0.02
Monroe (IN)	14	1	0.07
Monterey	25	2	0.08
Montgomery (AL)	41	3	0.07
Montgomery (OH)	231	3	0.01
Muscogee	80	4	0.05
Muskegon	20	0	0.00
Napa	9	0	0.00
Nashville	7	0	0.00
New Hanover	105	2	0.02
New Haven	374	56	0.15
New Orleans	353	35	0.10
New York (Greater)	1311	685	0.52
Newport News	77	3	0.04
Nueces	71	1	0.01
Oklahoma	470	36	0.08
Onondaga	181	12	0.07
Orange	527	52	0.10
Outagamie	5	0	0.00
Peoria	103	6	0.06
Philadelphia (Greater)	950	427	0.45
Phoenix	647	84	0.13

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Pima	269	23	0.09
Pitt	101	4	0.04
Pittsburgh	496	126	0.25
Polk	122	8	0.07
Portland	510	89	0.17
Potter	69	1	0.01
Providence	259	41	0.16
Pueblo	30	0	0.00
Pulaski	303	16	0.05
Punta Gorda	45	1	0.02
Racine	11	0	0.00
Richland	148	9	0.06
Richmond (Greater)	338	29	0.09
Roanoke	82	1	0.01
Rochester (MN)	240	120	0.50
Rochester (NY)	313	59	0.19
Rock	10	0	0.00
Sacramento	311	36	0.12
Saginaw	47	0	0.00
Salt Lake	439	73	0.17
San Antonio	652	105	0.16
San Diego	846	222	0.26
San Francisco (Greater)	822	368	0.45
San Joaquin	52	1	0.02
Sangamon	138	6	0.04
Santa Barbara	64	5	0.08
Santa Cruz	3	0	0.00
Sarasota	203	10	0.05
Scott	56	1	0.02
Seattle	705	216	0.31
Sebastian	23	0	0.00
Sedgwick	189	12	0.06
Shawnee	91	0	0.00
Sonoma	38	1	0.03
Spokane	238	9	0.04
St. Joseph	73	1	0.01
St. Louis	761	155	0.20
St. Lucie	31	2	0.06
Stanislaus	25	1	0.04
Stark	131	7	0.05
Summit	161	3	0.02
Sumter	4	0	0.00
Sutter	8	0	0.00
Tallahassee	50	1	0.02
Tampa-Hernando	28	0	0.00
Tampa-Hillsborough	447	58	0.13
Tampa-Pinellas	373	10	0.03
Taylor	10	1	0.10
Terrebonne	10	1	0.10
Thurston	48	2	0.04
Tulare	5	0	0.00
Tulsa	156	11	0.07
Tuscaloosa	38	2	0.05
Utah	54	0	0.00
Vanderburgh	99	2	0.02
Ventura	75	8	0.11
Virginia Beach	273	27	0.10
Volusia-Daytona Beach	213	4	0.02
Wake	231	18	0.08
Washington (Greater)	1058	433	0.41
Washington (MD)	39	0	0.00
Washoe	71	1	0.01
Washtenaw	293	91	0.31
Webb	8	0	0.00
Weber	60	1	0.02
Weld	18	0	0.00
Whatcom	43	1	0.02
Wichita	20	0	0.00
Winnebago (IL)	43	0	0.00
Winnebago (WI)	1	0	0.00
Woodbury	41	2	0.05
Worcester	151	13	0.09
Yakima	42	1	0.02
Yellowstone	103	2	0.02
York	20	1	0.05
Uruquay	**6**	**1**	**0.17**

(*continued on next page*)

(*continued*)

	Trial (n)	Author (n)	Rate (%)
Montevideo	6	1	0.17
Venezuala	**23**	**2**	**0.09**
Barquisimeto	1	0	0.00
Caracas	13	2	0.15
Ciudad Guayana	1	0	0.00
Maracaibo	5	0	0.00
Valencia	3	0	0.00
Vietnam	**38**	**4**	**0.11**
Hai Phong	1	0	0.00
Hanoi	18	3	0.17
Ho Chi Minh City	19	1	0.05
Zambia	**12**	**9**	**0.75**
Lusaka	12	9	0.75
Zimbabwe	**12**	**10**	**0.83**
Harare	12	10	0.83

References

Adams, J.D., Black, G.C., Clemmons, J.R., Stephan, P.E., 2005. Scientific teams and institutional collaborations: evidence from U.S. universities, 1981–1999. Res. Policy 34, 259–285. https://doi.org/10.1016/j.respol.2005.01.014.

Aliukonis, V., Poškutė, M., Gefenas, E., 2020. Perish or publish dilemma: challenges to responsible authorship. Medicina (Kaunas) 56, 123. https://doi.org/10.3390/medicina56030123.

Allen, L., Scott, J., Brand, A., Hlava, M., Altman, M., 2014. Publishing: credit where credit is due. Nature 508, 312–313. https://doi.org/10.1038/508312a.

Almeida, P., Kogut, B., 1999. Localization of knowledge and the mobility of engineers in regional networks. Manag. Sci. 45, 905–917. https://doi.org/10.1287/mnsc.45.7.905.

Archer, S.W., Carlo, W.A., Truog, W.E., Stevenson, D.K., Van Meurs, K.P., Sánchez, P.J., Das, A., Devaskar, U., Nelin, L.D., Petrie Huitema, C.M., Crawford, M.M., Higgins, R. D., 2016. Improving publication rates in a collaborative clinical trials research network. Semin. Perinatol. 40, 410–417. https://doi.org/10.1053/j.semperi.2016.05.003.

Awan, F.A., Becker, A.B., Wang, Y., Kimmelman, J., 2022. Participant recruitment from low- and middle-income countries for pivotal trials of drugs approved by the U.S. Food and Drug Administration. Ann. Intern. Med. 175, 1675–1684. https://doi.org/10.7326/M22-1857.

Azoulay, P., 2004. Capturing knowledge within and across firm boundaries: evidence from clinical development. Am. Econ. Rev. 94, 1591–1612. https://doi.org/10.1257/0002828043052259.

Azoulay, P., Repenning, N.P., Zuckerman, E.W., 2010. Nasty, brutish, and short: embeddedness failure in the pharmaceutical industry. Adm. Sci. Q. 55, 472–507. https://doi.org/10.2189/asqu.2010.55.3.472.

Beck, S., Bergenholtz, C., Bogers, M., Brasseur, T.-M., Conradsen, M.L., Di Marco, D., Distel, A.P., Dobusch, L., Dörler, D., Effert, A., Fecher, B., Filiou, D., Frederiksen, L., Gillier, T., Grimpe, C., Gruber, M., Haeussler, C., Heigl, F., Hoisl, K., Hyslop, K., Kokshagina, O., LaFlamme, M., Lawson, C., Lifshitz-Assaf, H., Lukas, W., Nordberg, M., Norn, M.T., Poetz, M., Ponti, M., Pruschak, G., Pujol Priego, L., Radziwon, A., Rafner, J., Romanova, G., Ruser, A., Sauermann, H., Shah, S.K., Sherson, J.F., Suess-Reyes, J., Tucci, C.L., Tuertscher, P., Vedel, J.B., Velden, T., Verganti, R., Wareham, J., Wiggins, A., Xu, S.M., 2022. The Open Innovation in Science research field: a collaborative conceptualisation approach. Ind. Innov. 29, 136–185. https://doi.org/10.1080/13662716.2020.1792274.

Biagioli, M., 1998. LIFE SCIENCES FORUM the instability of authorship: credit and responsibility in contemporary biomedicine. FASEB J. 12, 3–16. https://doi.org/10.1096/fsb2fasebj.12.1.3.

Biagioli, M., Galison, P., 2003. Scientific Authorship: Credit and Intellectual Property in Science. Psychology Press.

Bonaccorsi, A., 2010. New forms of complementarity in science. Minerva 48, 355–387. https://doi.org/10.1007/s11024-010-9159-6.

Bonitz, M., 2005. Ten years Matthew effect for countries. Scientometrics 64, 375–379. https://doi.org/10.1007/s11192-005-0256-5.

Boschma, R., Heimeriks, G., Balland, P.-A., 2014. Scientific knowledge dynamics and relatedness in biotech cities. Res. Policy 43, 107–114. https://doi.org/10.1016/j.respol.2013.07.009.

Breschi, S., Lissoni, F., 2009. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. J. Econ. Geogr. lbp008 https://doi.org/10.1093/jeg/lbp008.

Buchkowsky, S.S., Jewesson, P.J., 2004. Industry sponsorship and authorship of clinical trials over 20 years. Ann. Pharmacother. 38, 579–585. https://doi.org/10.1345/aph.1D267.

Catalini, C., Fons-Rosen, C., Gaulé, P., 2020. How do travel costs shape collaboration? Manag. Sci. 66, 3340–3360. https://doi.org/10.1287/mnsc.2019.3381.

Ciarli, T., Ràfols, I., 2019. The relation between research priorities and societal demands: the case of rice. In: Research Policy, New Frontiers in Science, Technology and Innovation Research From SPRU's 50th Anniversary Conference, 48, pp. 949–967. https://doi.org/10.1016/j.respol.2018.10.027.

Cockburn, I.M., Henderson, R.M., 1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. J. Ind. Econ. 46, 157–182. https://doi.org/10.1111/1467-6451.00067.

Collins, H., 2010. Tacit and Explicit Knowledge. University of Chicago Press.

Confraria, H., Wang, L., 2020. Medical research versus disease burden in Africa. Res. Policy 49, 103916. https://doi.org/10.1016/j.respol.2019.103916.

Csomós, G., 2018. A spatial scientometric analysis of the publication output of cities worldwide. J. Informet. 12, 547–566. https://doi.org/10.1016/j.joi.2018.05.003.

Dasgupta, P., David, P.A., 1994. Toward a new economics of science. Res. Policy 23, 487–521. https://doi.org/10.1016/0048-7333(94)01002-1.

Dijkstra, L., Poelman, H., Veneri, P., 2019. The EU-OECD definition of a functional urban area. In: OECD Regional Development Working Papers, No. 2019/11. OECD Publishing, Paris. https://doi.org/10.1787/d58cb34d-en.

Dombernowsky, T., Haedersdal, M., Lassen, U., Thomsen, S.F., 2019. Criteria for site selection in industry-sponsored clinical trials: a survey among decision-makers in biopharmaceutical companies and clinical research organizations. Trials 20, 708. https://doi.org/10.1186/s13063-019-3790-9.

Dulhunty, J.M., Boots, R.J., Paratz, J.D., Lipman, J., 2011. Determining authorship in multicenter trials: a systematic review. Acta Anaesthesiol. Scand. 55, 1037–1043. https://doi.org/10.1111/j.1399-6576.2011.02477.x.

Fisher, J.A., 2008. Medical Research for Hire: The Political Economy of Pharmaceutical Clinical Trials. Rutgers University Press.

Flanagin, A., Carey, L.A., Fontanarosa, P.B., Phillips, S.G., Pace, B.P., Lundberg, G.D., Rennie, D., 1998. Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. JAMA 280, 222–224. https://doi.org/10.1001/jama.280.3.222.

Frenken, K., Heimeriks, G.J., Hoekman, J., 2017. What drives university research performance? An analysis using the CWTS Leiden Ranking data. J. Informet. 11, 859–872. https://doi.org/10.1016/j.joi.2017.06.006.

Gasparyan, A.Y., Ayvazyan, L., Kitas, G.D., 2013. Authorship problems in scholarly journals: considerations for authors, peer reviewers and editors. Rheumatol. Int. 33, 277–284. https://doi.org/10.1007/s00296-012-2582-2.

Gehring, M., Taylor, R.S., Mellody, M., Casteels, B., Piazzi, A., Gensini, G., Ambrosio, G., 2013. Factors influencing clinical trial site selection in Europe: the Survey of Attitudes towards Trial sites in Europe (the SAT-EU Study). BMJ Open 3, e002957. https://doi.org/10.1136/bmjopen-2013-002957.

Gelijns, A., Rosenberg, N., 1994. The dynamics of technological change in medicine. Health Aff. 13, 28–46. https://doi.org/10.1377/hlthaff.13.3.28.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M., 1994. The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies. SAGE Publications.

Gittelman, M., 2016. The revolution re-visited: clinical and genetics research paradigms and the productivity paradox in drug discovery. Res. Policy. https://doi.org/10.1016/j.respol.2016.01.007.

Glickman, S.W., McHutchison, J.G., Peterson, E.D., Cairns, C.B., Harrington, R.A., Califf, R.M., Schulman, K.A., 2009. Ethical and scientific implications of the globalization of clinical research. N. Engl. J. Med. 360, 816–823. https://doi.org/10.1056/NEJMsb0803929.

Haeussler, C., Rake, B., 2017. The changing geography of clinical research: a critical analysis of its drivers. Ind. Corp. Chang. 26, 285–310. https://doi.org/10.1093/icc/dtx002.

Haeussler, C., Sauermann, H., 2013. Credit where credit is due? The impact of project contributions and social factors on authorship and inventorship. Res. Policy 42 (3), 688–703.

Haeussler, C., Sauermann, H., 2020. Division of labor in collaborative knowledge production: the role of team size and interdisciplinarity. Res. Policy 49, 103987. https://doi.org/10.1016/j.respol.2020.103987.

Hall, K.L., Vogel, A.L., Huang, G.C., Serrano, K.J., Rice, E.L., Tsakraklides, S.P., Fiore, S. M., 2018. The science of team science: a review of the empirical evidence and

research gaps on collaboration in science. Am. Psychol. 73, 532–548. https://doi.org/10.1037/amp0000319.
Hawkins, B.S., 2020. Publications from clinical trials. In: Piantadosi, S., Meinert, C.L. (Eds.), Principles and Practice of Clinical Trials. Springer International Publishing, Cham, pp. 1–14. https://doi.org/10.1007/978-3-319-52677-5_184-1.
Heimeriks, G., Boschma, R., 2013. The path- and place-dependent nature of scientific knowledge production in biotech 1986–2008. J. Econ. Geogr. lbs052 https://doi.org/10.1093/jeg/lbs052.
Hessels, L.K., van Lente, H., Smits, R., 2009. In search of relevance: the changing contract between science and society. Sci. Public Policy 36, 387–401. https://doi.org/10.3152/030234209X442034.
Hoekman, J., Frenken, K., Oort, F. van, 2009. The geography of collaborative knowledge production in Europe. Ann. Reg. Sci. 43, 721–738. https://doi.org/10.1007/s00168-008-0252-9.
Hoekman, J., Frenken, K., Tijssen, R.J.W., 2010. Research collaboration at a distance: changing spatial patterns of scientific collaboration within Europe. Res. Policy 39, 662–673. https://doi.org/10.1016/j.respol.2010.01.012.
Hoekman, J., Frenken, K., de Zeeuw, D., Heerspink, H.L., 2012. The geographical distribution of leadership in globalized clinical trials. PLoS One 7, e45984. https://doi.org/10.1371/journal.pone.0045984.
Holcombe, A.O., 2019. Contributorship, not authorship: use credit to indicate who did what. Publications 7, 48. https://doi.org/10.3390/publications7030048.
Hong, W., 2008. Domination in a scientific field: capital struggle in a chinese isotope lab. Soc. Stud. Sci. 38, 543–570. https://doi.org/10.1177/0306312706092456.
Jabbehdari, S., Walsh, J.P., 2017. Authorship norms and project structures in science. Sci. Technol. Hum. Values 42, 872–900. https://doi.org/10.1177/0162243917697192.
Jones, B.F., Wuchty, S., Uzzi, B., 2008. Multi-university research teams: shifting impact, geography, and stratification in science. Science 322, 1259–1262. https://doi.org/10.1126/science.1158357.
Kassirer, J.P., Angell, M., 1991. On authorship and acknowledgments. N. Engl. J. Med. 325, 1510–1512. https://doi.org/10.1056/NEJM199111213252112.
Katz, J.S., 1994. Geographical proximity and scientific collaboration. Scientometrics 31, 31–43. https://doi.org/10.1007/BF02018100.
Katz, J.S., Martin, B.R., 1997. What is research collaboration? Res. Policy 26, 1–18. https://doi.org/10.1016/S0048-7333(96)00917-1.
Kaufmann, P., Annis, C., Griggs, R.C., 2010. The authorship lottery: an impediment to research collaboration? Ann. Neurol. 68, 782–786. https://doi.org/10.1002/ana.22232.
Kelaher, M., Ng, L., Knight, K., Rahadi, A., 2016. Equity in global health research in the new millennium: trends in first-authorship for randomized controlled trials among low- and middle-income country researchers 1990-2013. Int. J. Epidemiol. 45, 2174–2183. https://doi.org/10.1093/ije/dyw313.
Lane, J., 2010. Let's make science metrics more scientific. Nature 464, 488–489. https://doi.org/10.1038/464488a.
Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., Sugimoto, C.R., 2016. Contributorship and division of labor in knowledge production. Soc. Stud. Sci. 46, 417–435. https://doi.org/10.1177/0306312716650046.
Latour, B., Woolgar, S., 1979. Laboratory Life: The Social Construction of Scientific Facts.
Laudel, G., 2002. What do we measure by co-authorships? Res. Eval. 11, 3–15. https://doi.org/10.3152/147154402781776961.
Lissoni, F., Montobbio, F., Zirulia, L., 2013. Inventorship and authorship as attribution rights: an enquiry into the economics of scientific credit. J. Econ. Behav. Organ. 95, 49–69. https://doi.org/10.1016/j.jebo.2013.08.016.
Lissoni, F., Montobbio, F., Zirulia, L., 2020. Misallocation of scientific credit: the role of hierarchy and preferences. An extension of Lissoni et al. (2013). Ind. Corp. Chang. 29, 1471–1482. https://doi.org/10.1093/icc/dtaa029.
Malmberg, A., Maskell, P., 2002. The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering. Environ. Plan. A 34, 429–449. https://doi.org/10.1068/a3457.
Marschke, G., Nunez, A., Weinberg, B.A., Yu, H., 2018. Last place? The intersection of ethnicity, gender, and race in biomedical authorship. AEA Pap. Proc. 108, 222–227. https://doi.org/10.1257/pandp.20181111.
Marshall, I.J., L'Esperance, V., Marshall, R., Thomas, J., Noel-Storr, A., Soboczenski, F., Nye, B., Nenkova, A., Wallace, B.C., 2021. State of the evidence: a survey of global disparities in clinical trials. BMJ Glob. Health 6, e004145. https://doi.org/10.1136/bmjgh-2020-004145.
Marušić, A., Bošnjak, L., Jerončić, A., 2011. A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. PLoS One 6, e23477. https://doi.org/10.1371/journal.pone.0023477.
McNutt, M.K., Bradford, M., Drazen, J.M., Hanson, B., Howard, B., Jamieson, K.H., Kiermer, V., Marcus, E., Pope, B.K., Schekman, R., Swaminathan, S., Stang, P.J., Verma, I.M., 2018. Transparency in authors' contributions and responsibilities to promote integrity in scientific publication. Proc. Natl. Acad. Sci. 115, 2557–2560. https://doi.org/10.1073/pnas.1715374115.
Mentz, R.J., Peterson, E.D., 2017. Site principal investigators in multicenter clinical trials: appropriately recognizing key contributors. Circulation 135, 1185–1187. https://doi.org/10.1161/CIRCULATIONAHA.116.026650.
Merton, R.K., 1973. The Sociology of Science: Theoretical and Empirical Investigations. University of Chicago Press.
Metcalfe, J.S., James, A., Mina, A., 2005. Emergent innovation systems and the delivery of clinical services: the case of intra-ocular lenses. Res. Policy 34, 1283–1304. https://doi.org/10.1016/j.respol.2005.01.015.
Miller, J., Millum, J., 2022. Ethical considerations in international clinical trial site selection. BMJ Glob. Health 7, e008012. https://doi.org/10.1136/bmjgh-2021-008012.
Milojević, S., 2014. Principles of scientific research team formation and evolution. PNAS 111, 3984–3989. https://doi.org/10.1073/pnas.1309723111.
Mina, A., Ramlogan, R., Tampubolon, G., Metcalfe, J.S., 2007. Mapping evolutionary trajectories: applications to the growth and transformation of medical knowledge. Res. Policy 36, 789–806. https://doi.org/10.1016/j.respol.2006.12.007.
Morton, B., Vercueil, A., Masekela, R., Heinz, E., Reimer, L., Saleh, S., Kalinga, C., Seekles, M., Biccard, B., Chakaya, J., Abimbola, S., Obasi, A., Oriyo, N., 2022. Consensus statement on measures to promote equitable authorship in the publication of research from international partnerships. Anaesthesia 77, 264–276. https://doi.org/10.1111/anae.15597.
Moynihan, R., 2008. Key opinion leaders: independent experts or drug representatives in disguise? BMJ 336, 1402–1403. https://doi.org/10.1136/bmj.39575.675787.651.
Murray, C.J.L., Lopez, A.D., Organization, W.H, Bank, W, Health, H.S. of P, 1996. The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability From Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020: Summary. World Health Organization.
Nomaler, Ö., Frenken, K., Heimeriks, G., 2013. Do more distant collaborations have more citation impact? J. Informet. 7, 966–971. https://doi.org/10.1016/j.joi.2013.10.001.
Nomaler, Ö., Frenken, K., Heimeriks, G., 2014. On scaling of scientific knowledge production in U.S. metropolitan areas. PLoS One 9, e110805. https://doi.org/10.1371/journal.pone.0110805.
Obasi, A.I.N., Abimbola, S., Oriyo, N., Morton, B., Vercueil, A., Masekela, R., 2021. Credit local authors fairly on international research papers. Nature 598, 415. https://doi.org/10.1038/d41586-021-02831-w.
Pan, R.K., Kaski, K., Fortunato, S., 2012. World citation and collaboration networks: uncovering the role of geography in science. Sci. Rep. 2 https://doi.org/10.1038/srep00902.
Patterson, T., 2012. Creating natural earth map data. http://www.shadedrelief.com/new_zealand/natural_earth_nzcs.pdf. (Accessed 23 October 2023).
Penders, B., 2016. Letter to the editor: respecting the plurality of value and the messiness of scientific practice. Account. Res. 23, 136–138. https://doi.org/10.1080/08989621.2015.1060128.
Petryna, A., 2009. When Experiments Travel: Clinical Trials and the Global Search for Human Subjects. Princeton University Press.
Polanyi, M., 1966. The Tacit Dimension. Doubleday, Garden City, N.Y.
Polidoro, F., Theeke, M., 2012. Getting competition down to a science: the effects of technological competition on firms' scientific publications. Organ. Sci. 23, 1135–1153. https://doi.org/10.1287/orsc.1110.0684.
Ponds, R., Oort, F.V., Frenken, K., 2007. The geographical and institutional proximity of research collaboration*. Pap. Reg. Sci. 86, 423–443. https://doi.org/10.1111/j.1435-5957.2007.00126.x.
Qiu, S., Steinwender, C., Azoulay, P., 2022. Who Stands on the Shoulders of Chinese (Scientific) Giants? Evidence From Chemistry. NBER Working Paper, 30772. https://doi.org/10.3386/w30772.
Rafols, I., Hopkins, M.M., Hoekman, J., Siepel, J., O'Hare, A., Perianes-Rodríguez, A., Nightingale, P., 2013. Big Pharma, little science?: A bibliometric perspective on Big Pharma's R&D decline. Technol. Forecast. Soc. Chang. https://doi.org/10.1016/j.techfore.2012.06.007.
Rasmussen, N., 2005. The Drug Industry and Clinical Research in Interwar America: three types of physician collaborator. Bull. Hist. Med. 79, 50–80.
Rasmussen, K., Bero, L., Redberg, R., Gøtzsche, P.C., Lundh, A., 2018. Collaboration between academics and industry in clinical trials: cross sectional study of publications and survey of lead academic authors. BMJ 363. https://doi.org/10.1136/bmj.k3654.
Rees, C.A., Keating, E.M., Dearden, K.A., Haq, H., Robison, J.A., Kazembe, P.N., Bourgeois, F.T., Niescierenko, M., 2019. Importance of authorship and inappropriate authorship assignment in paediatric research in low- and middle-income countries. Tropical Med. Int. Health 24, 1229–1242. https://doi.org/10.1111/tmi.13295.
Rennie, D., Yank, V., Emanuel, L., 1997. When authorship fails: a proposal to make contributors accountable. JAMA 278, 579–585. https://doi.org/10.1001/jama.1997.03550070071041.
Rettig, R.A., 2000. The industrialization of clinical research. Health Aff. 19, 129–146. https://doi.org/10.1377/hlthaff.19.2.129.
Rosenberg, J., Burcharth, J., Pommergaard, H.-C., Vinther, S., 2015. Authorship issues in multi-centre clinical trials: the importance of making an authorship contract. Dan. Med. J. 62.
Ross, M.B., Glennon, B.M., Murciano-Goroff, R., Berkes, E.G., Weinberg, B.A., Lane, J.I., 2022. Women are credited less in science than men. Nature 608, 135–145. https://doi.org/10.1038/s41586-022-04966-w.
Sauermann, H., Haeussler, C., 2017. Authorship and contribution disclosures. Sci. Adv. 3, e1700404 https://doi.org/10.1126/sciadv.1700404.
Sauermann, H., Stephan, P., 2013. Conflicting logics? A multidimensional view of industrial and academic science. Organ. Sci. 24, 889–909. https://doi.org/10.1287/orsc.1120.0769.
Simcoe, T.S., Waguespack, D.M., 2010. Status, quality, and attention: what's in a (missing) name? Manag. Sci. 57, 274–290. https://doi.org/10.1287/mnsc.1100.1270.
Sismondo, S., 2009. Ghosts in the machine: publication planning in the medical sciences. Soc. Stud. Sci. 39, 171–198. https://doi.org/10.1177/0306312708101047.
Smith, E., 2023. "Technical" contributors and authorship distribution in health science. Sci. Eng. Ethics 29, 22. https://doi.org/10.1007/s11948-023-00445-1.
Smith, E., Hunt, M., Master, Z., 2014. Authorship ethics in global health research partnerships between researchers from low or middle income countries and high income countries. BMC Med. Ethics 15, 42. https://doi.org/10.1186/1472-6939-15-42.

Smith, E., Williams-Jones, B., Master, Z., Larivière, V., Sugimoto, C.R., Paul-Hus, A., Shi, M., Diller, E., Caudle, K., Resnik, D.B., 2020. Researchers' perceptions of ethical authorship distribution in collaborative research teams. Sci. Eng. Ethics 26, 1995–2022. https://doi.org/10.1007/s11948-019-00113-3.

Stephan, P., 2012. How Economics Shapes Science.

Tauber, M., Paul, C., 2017. Authorship selection in industry-sponsored publications of dermatology clinical trials. Br. J. Dermatol. 176, 1669–1671. https://doi.org/10.1111/bjd.15459.

Thiers, F.A., Sinskey, A.J., Berndt, E.R., 2008. Trends in the globalization of clinical trials. Nat. Rev. Drug Discov. 7, 13–14. https://doi.org/10.1038/nrd2441.

Wager, E., 2009. Recognition, reward and responsibility: why the authorship of scientific papers matters. Maturitas 62, 109–112. https://doi.org/10.1016/j.maturitas.2008.12.001.

Walsh, J.P., Lee, Y.-N., 2015. The bureaucratization of science. Res. Policy 44, 1584–1600. https://doi.org/10.1016/j.respol.2015.04.010.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., Eck, N.J. van, Leeuwen, T.N. van, Raan, A.F.J. van, Visser, M.S., Wouters, P., 2012. The Leiden ranking 2011/2012: data collection, indicators, and interpretation. J. Am. Soc. Inf. Sci. Technol. 63, 2419–2432. https://doi.org/10.1002/asi.22708.

Wang, J., Hicks, D., 2015. Scientific teams: self-assembly, fluidness, and interdependence. J. Inform. 9 (1), 197–207.

Whellan, D.J., Ellis, S.J., Kraus, W.E., Hawthorne, K., Piña, I.L., Keteyian, S.J., Kitzman, D.W., Cooper, L., Lee, K., O'Connor, C.M., 2009. Method for establishing authorship in a multicenter clinical trial. Ann. Intern. Med. 151, 414–420. https://doi.org/10.7326/0003-4819-151-6-200909150-00006.

Whellan, D.J., Kraus, W.E., Kitzman, D.W., Rooney, B., Keteyian, S.J., Piña, I.L., Ellis, S. J., Ghali, J.K., Lee, K.L., Cooper, L.S., O'Connor, C.M., 2015. Authorship in a multicenter clinical trial: The Heart Failure-A Controlled Trial Investigating Outcomes of Exercise Training (HF-ACTION) Authorship and Publication (HAP) scoring system results. Am. Heart J. 169, 457–463.e6. https://doi.org/10.1016/j.ahj.2014.11.022.

Wislar, J.S., Flanagin, A., Fontanarosa, P.B., DeAngelis, C.D., 2011. Honorary and ghost authorship in high impact biomedical journals: a cross sectional survey. BMJ 343, d6128. https://doi.org/10.1136/bmj.d6128.

World Health Organisation, 2018. WHO methods and data sources for global burden of disease estimates 2000-2016. In: Global Health Estimates Technical Paper WHO/HIS/IER/GHE/2018.4.

Wouters, P., 1997. Citation cycles and peer review cycles. Scientometrics 38, 39–55. https://doi.org/10.1007/BF02461122.

Wu, L., Wang, D., Evans, J.A., 2019. Large teams develop and small teams disrupt science and technology. Nature 566, 378–382. https://doi.org/10.1038/s41586-019-0941-9.

Wuchty, S., Jones, B.F., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. Science 316, 1036–1039. https://doi.org/10.1126/science.1136099.

Wuestman, M.L., Hoekman, J., Frenken, K., 2019. The geography of scientific citations. Res. Policy 48, 1771–1780. https://doi.org/10.1016/j.respol.2019.04.004.

Yank, V., Rennie, D., 1999. Disclosure of researcher contributions: a study of original research articles in the Lancet. Ann. Intern. Med. 130, 661–670. https://doi.org/10.7326/0003-4819-130-8-199904200-00013.

Zuckerman, H.A., 1968. Patterns of name ordering among authors of scientific papers: a study of social symbolism and its ambiguity. Am. J. Sociol. 74, 276–291. https://doi.org/10.1086/224641.