If You Build It, They Will Come: The Impact of Clinical Trial Experience on African Science

Caroline Fry[1] and Michael Blomfield[2]

[1]Shidler College of Business, University of Hawai'i at Manoa
[2]University of Massachusetts, Amherst

April 10, 2024

Abstract

Advancing innovative solutions to address significant global health challenges often requires individuals who possess the necessary knowledge and skills to execute critical tasks. One possible way to shape scientists' research trajectories so that they develop the skills to undertake these tasks is through the provision of experience. This paper investigates how participation as a researcher in a clinical trial influences the involvement of scientists in Sub-Saharan Africa in subsequent, similar projects. We document that scientists who participate on the research team in a clinical trial are more likely to be involved in future trials, particularly in the same disease area, with the most significant impact seen in scientists without prior trial experience and for earlier career scientists. This suggests that knowledge and skills acquired in the initial project are transferable to subsequent endeavors. However, we also find that scientists, and particularly more experienced scientists, working in countries with lower initial trial capacity have limited increases in subsequent clinical trial involvement following trial experience. The results suggest that efforts to improve innovative capabilities to solve health challenges will require both investments in individuals to provide the necessary experience and efforts to improve the institutional environments in which the researchers are embedded.

1 Introduction

In 2021, the world's first malaria vaccine, Mosquirix, was approved for global use by the World Health Organization. Development of this vaccine started in 1987, with Phase III clinical trials taking place from 2003 onwards across several African countries. Overall, the vaccine spent 23 years in trials and pilot studies prior to being licensed. Two years later, in 2023, a second malaria vaccine, Matrix-M, was approved for regulatory use in Ghana. This time, the African-led Phase III trials for this vaccine started just two years prior to vaccine approval. In speaking about his team's role in the latter vaccine trials, Professor Abdoulaye Djimde in Mali said: *'We will utilize our more than two decades of experience in malaria vaccine testing towards successful completion of this trial.'*[1] This example raises the question: How, and under what circumstances, does the opportunity to gain experience in an innovative process shape a researcher's future research trajectory?

Developing effective treatments for high impact diseases such as malaria is an archetype of the global grand challenges that can only be solved by coordinated action by a wide range of stakeholders (Arslan and Tarakci, 2022; George et al., 2016; Howard-Grenville, 2021; George et al., 2024). Past research on grand challenges has typically focused primarily on questions of governance across the range of actors involved (Couture et al., 2023; Dentoni et al., 2018), how institutional reforms may incentivize actors to focus more on grand challenges (Vakili and McGahan, 2016), or organizations' prioritization of 'responsible innovation' (Owen et al., 2021). Other research has shown how changes in policy institutions, such as Intellectual Property regimes, can also facilitate greater scientific research and diffusion of knowledge on neglected diseases (Vakili and McGahan, 2016). However, far less research has focused on the scientists themselves who are responsible for advancing solutions to grand challenges. Consider the research and development of vaccines and drugs for neglected diseases, for example. The innovation process underlying drug development requires scientists with relevant capabilities to undertake essential tasks, from basic research to conducting large-scale clinical trials within affected populations. Given the central role of individuals with skills, knowledge and preferences to study relevant topics in solving grand challenges, a better understanding of how scientists' research trajectories are shaped in this context is critical.

In this paper we highlight the role of experience in forming scientists' research trajectories. Engaging in specific innovative processes can foster 'learning by doing,'[2] leading to the acquisition of specialized knowledge and skills. Consequently, scientists can focus on research types that require these skills, thereby shaping their research trajectory. Beyond skills, scientists could develop domain-specific infrastructure, networks and could gain superior access to relevant opportunities as a result of an experience. This may also shape their research trajectory because it makes participation in similar projects more accessible or lower cost. In line with this argument, Azoulay et al. (2009) show that scientists tend to produce more

[1] https://www.ox.ac.uk/news/2021-05-07-promising-malaria-vaccine-enters-final-stage-clinical-testing-west-africa last accessed on 1.31.24

[2] For a review of the literature on learning by doing see (Argote et al., 2021)

commercially oriented research following patenting of one of their ideas. The authors propose that these scientists develop new relationships with industry-based researchers after a patenting experience which exposes them to new questions, shaping the content of their research. That said, it is not immediately obvious that scientists would be able, or willing, to select projects in line with their past experience. Project selection is the culmination of a number of factors, ranging from existing incentives, financial reward available, competition in ideas space and the local environment and policies (see Teodoridis et al. (2022) for a review of the literature on factors driving scientists' project selection). In particular, scientists in environments with fewer resources may lack opportunities to selectively participate in innovative processes in specific domains0. This potentially constrains their capacity to accumulate highly specialized capabilities and ultimately limiting the influence of experience on their research trajectory.

We explore the role of experience in scientists' subsequent research trajectories in the context of clinical trials in Sub-Saharan Africa. This is an important context to study these questions. More efficient testing and development of potential drugs, vaccines, and treatment regimens that target diseases such as malaria, TB and HIV that have a large burden in Sub-Saharan Africa, in particular, would have enormous impact on the lives of those affected by these diseases. Yet, despite the disproportionate burden of global disease faced by Sub-Saharan Africa, relatively few clinical trials take place in the region,[3] and those that do take place are predominantly supported by foreign funders.[4]. More broadly, there is growing recognition of the importance of having diverse populations participate in clinical trials to ensure efficacy among the intended recipient groups (Arslan and Tarakci, 2022). One major barrier to increasing the number of clinical trials in Sub-Saharan Africa is a lack of experienced scientists and health professionals with the knowledge and skills to carry out clinical trials (Alemayehu et al., 2018). While experience in the research team in a clinical trial could present an opportunity for scientists to develop specific capabilities, positively impacting their ability to participate in future similar clinical trials, there could be limits to the extent to which scientists can leverage their experience. Namely, given a high reliance on attracting external resources in this setting, access to opportunities to participate in subsequent clinical trials could be limited. For example, scientists' countries may lack transparent regulatory systems with established procedures to facilitate trials (Taylor-Robinson et al., 2021), making it harder for scientists to attract external sponsors.

We conduct our analysis using data on clinical trial participation of African scientists. The European and Developing Countries Clinical Trial Partnership (EDCTP) is an EU-funded partnership between European and Sub-Saharan African countries that has funded more than 300 clinical trials in Africa to date, and are the primary funder of clinical trials in Africa. These are primarily focused on HIV, malaria, and tuberculosis and are mostly at Phases II and III bringing together scientists from European and African countries to work on a clinical trial. The EDCTP offers funding at both the trial level and for individual scientists' career development by sponsoring their participation in a trial. We evaluate the relationship between par-

[3]The overwhelming majority of clinical trials take place in the United States and Europe. According to a recent assessment, only 2.5% of trials take place in Sub-Saharan Africa (Taylor-Robinson et al., 2021)

[4]Funding for research in Africa is very low, with governments committing just 0.42% of GDP in 2019, compared to the global average of 1.7% (Adepoju, 2022)

ticipating in an EDCTP funded trial and subsequent trial participation and research trajectory of scientists in Sub-Saharan Africa. To do so, we construct a panel dataset of 880 African scientists who participated in an EDCTP trial between 2003 and 2015 and had an observable publication record prior to the trial. For our analysis, we combine data on clinical trial involvement from ClinicalTrials.gov with publication data. We complement the quantitative data with qualitative data gathered during interviews with clinical trial funders, scientists and regulatory and coordinating bodies. We match scientists who participated in an EDCTP trial with a control group of scientists who are those who did not participate in an EDCTP trial but were otherwise similar in terms of their career age, prior international collaborations, extent of previous clinical trial and applied research, and their institution's size and prior trial involvement.

We use difference-in-differences regressions to compare the within-scientist changes in scientists' clinical trial involvement and research output after they participate in an EDCTP-funded trial with the changes of observably similar scientists who were not involved in an EDCTP-funded trial. We find that participation in an ECDTP-funded trial is associated with scientists' having greater subsequent involvement in clinical trials and shifting their research to focus more on applied, disease-oriented research, particularly in their trial's disease. Treated scientists undertake future clinical trials at approximately three times the rate of control scientists after participating in an EDCTP trial. This is an economically significant increase that reflects a 300 percent increase relative to the mean rate of clinical trials. Interestingly, they do not increase the rate at which they participate in African sponsored trials,[5] but they do tend to be involved in both more EU and US sponsored trials, implying that attracting external resources is critical in this context.

To understand how and why experience in trials is associated with scientists' future research trajectory, we unpack the heterogeneity in outcomes among those scientists who participate in the EDCTP program. This allows us to focus on how individual and national level factors affect how scientists leverage trial experience in their subsequent research. At the individual-level, we find that the increase in clinical trial involvement is most pronounced for scientists without prior participation in clinical trials. In fact, scientists who have previously worked on clinical trials show little evidence of increased participation in future trials relative to matched control scientists with similar prior clinical trial experience. We find that the trial-linked publications of scientists without prior trial experience include a greater share of new keywords that have not previously been indexed to the scientist's research, compared to those with prior trial experience. This is in line with an interpretation in which working on a first clinical trial is important for scientists to build the knowledge and skills that facilitates participation in future trials, with diminished returns to repeat participation.

Finally, we document that scientists in countries with higher levels of prior trials are more likely to participate in subsequent clinical trials compared to those in countries with fewer trials. Moreover, the difference in change in trial participation between scientists with and without prior experience is more salient in countries with fewer trials overall . This finding implies that the availability of trial opportunities in

[5]African sponsors accounted for just 3% of sponsor of clinical trials in our sample.

constrained environments significantly influences scientists', and particularly more experienced scientists', ability to leverage clinical trial experience to shift their research trajectory.

While there are limitations to making causal claims in the study, the paper does provide insight into how the trajectory of scientists' research is shaped. With this, we aim to make several contributions. First, we contribute to literature on the determinants of research direction. While prior research in this area has explored the role of funding availability (Myers, 2020), exposure to problems (Truffa and Wong, 2022; Fry, 2023), peers (Catalini, 2018; Azoulay et al., 2019), and access to data and tools (Nagaraj et al., 2020; Furman and Teodoridis, 2020), in this paper we highlight the role of experience in shaping the research trajectory of scientists. Namely, we demonstrate the path dependency of science and provide evidence that participation in a single project can, in some instances, substantially alter the long-term orientation of a scientist's agenda. Closest to our study is (Azoulay et al., 2009) who demonstrate that scientists tend to orient their research toward more commercializable projects following the patenting of an idea. We extend this work by showing that project experience can determine a scientist's long-run research trajectory and that learning curves can be surmounted relatively quickly, but that there are limits to the influence of an experience on future research. Namely, we show that the wider environment affects research direction, and that path dependency is not a given for scientists with environmental constraints. This implies that organizations seeking to foster individuals with specific capabilities can effectively increase this through providing experiences to a more diverse set of scientists, but that they might also consider investments in their supporting infrastructure and provision of follow on opportunities to nurture these new skills.

Second, we contribute to a growing literature on the role of formative experiences on scientists' careers. Prior research has documented the impact of researchers' human capital investments, or efforts to develop knowledge and skills through education, training, and experience, on their individual and organizational performance (Fong Boh et al., 2007; Jain, 2013; Roche, 2023; Levin and Stephan, 1991; Stephan, 2012; Shibayama, 2019; Conti et al., 2013; Kaiser et al., 2018; Zwick et al., 2017; Azoulay et al., 2017, 2009), resource acquisition (Kolympiris et al., 2019; Hoenen and Kolympiris, 2019; Stephan, 2012) and career choices (Gambardella et al., 2015; Azoulay et al., 2021; Agarwal and Ohyama, 2013; Sauermann and Roach, 2012; Ginther and Heggeness, 2020). We show that project experience is an important factor in ensuring that scientists have the necessary capabilities to increase their engagement in similar innovative projects, in this case, clinical trials.

Third, we contribute to research on the role of national institutions in shaping innovation (Fry and Furman, 2023; Vasudeva et al., 2013; Wang, 2015) by demonstrating the importance of national innovative capacity in determining scientists' ability to apply any benefits gained from research experiences.

More broadly, we respond to calls for researchers in the strategy and management fields to contribute insights from our disciplines to help find solutions to global grand challenges, and, in particular, to those in global health (George et al., 2016; Howard-Grenville, 2021; George et al., 2024; Arslan and Tarakci, 2022). We contribute to this literature by highlighting the important role of the development of the knowledge

and skills amongst the people carrying out research to solve grand challenges.

2 Theoretical Development

Innovation to solve global grand challenges requires high skilled human capital, or individuals who possess relevant knowledge, skills and preferences, to work on these problems. However, research skills and domain specific knowledge are challenging and costly to develop, requiring time actually doing the task to become proficient and often exchange with experts. Even if researchers develop the relevant skills and knowledge, directing scientists towards solving specific problems raises its own set of challenges. Extant empirical evidence demonstrates that researchers' project selection is influenced by a number of factors, including reputational concerns (Fry et al., 2023; Franzoni et al., 2011), peers and competition in ideas space (Catalini, 2018; Azoulay et al., 2019), exposure to problems (Truffa and Wong, 2022; Fry, 2023), finances available (Myers, 2020) and access to data and tools (Nagaraj et al., 2020; Furman and Teodoridis, 2020), and that re-directing scientists efforts is costly (Myers, 2020), and requires significant complementary assets to be accessible to individual scientists (Fry, 2023).

Less attention in this literature has been devoted to exploring the connection between the development of knowledge and skills and a researcher's orientation-and, in particular, the possibility that a single project experience could shape a scientist's research trajectory. Providing scientists opportunities to work on projects in which they can develop additional expertise may be an effective tool for policymakers to re-direct the scientists' efforts towards particular types of research. In this paper, we examine how, and under which circumstances, the opportunity to gain experience in an innovative process shapes a researcher's future research trajectory.

2.1 (When) can experience shape scientists' research trajectories?

Becker (1962) established the importance of human capitall—people's knowledge and skills—to their ability to carry out productive activities. Across a wide-range of tasks, individuals, groups, and organizations become more productive as their stock of experience with that type of task increases (Argote, 2012; Argote and Miron-Spektor, 2011). Early research in psychology established that learning curves exist at an individual level, with completion times and error rates on tasks decreasing in an individual's task experience (Ebbinghaus, 1885; Thurstone, 1919). The basic idea is that investing in accumulating knowledge and skills enables more effective work on the set of tasks for which this human capital is relevant. Learning through experience and working with peers can be especially important where knowledge or skills involve a tacit component that cannot be easily learned from codified sources (Polanyi, 1966; Raelin, 1997; Chan et al., 2014). Learning curves are scaled as individuals' expertise increases in their experience of task, or set of tasks, increasing the quality of their performance (Chase and Simon, 1973; Ericsson et al., 1993), and their understanding of where and when to apply those skills (Simon, 1991; Dane, 2010; Greenwood et al., 2019).

Prior research has documented the importance of learning curves for performance in a wide range of settings. These include manufacturing (Argote et al., 1990; Adler and Clark, 1991; Epple et al., 1991; Thompson, 2007) and service industries (Darr et al., 1995), alongside highly knowledge intensive settings such as medicine (Pisano et al., 2001; Reagans et al., 2005), software development (Fong Boh et al., 2007), and biotechnology research (Jain, 2013). One relatively understudied context in which learning curves are likely to be particularly important is amongst scientists. As science advances, the stock of knowledge that scientists require to carry out research at the frontier of their field becomes ever greater (Jones, 2009). This leads to scientists developing a narrower, more specialized research focus and increases the amount of time they must spend engaged in learning in order to develop needed knowledge and skills (Jones, 2009; Stephan, 2012; Agrawal et al., 2016). For scientists, accumulated prior research experience in a particular domain can increase productivity in that domain by providing a larger stock of knowledge and skills upon which they can draw. Laboratory-based training at undergraduate, graduate and post-doctoral levels facilitate scientists' acquisition and refinement of skills through carrying out research activities with experienced scientists (Ravetz, 1971). Beyond the development of knowledge and skills, accumulated experience could result in the development of domain-specific infrastructure, and superior access to networks and relevant opportunities, enhancing productivity and lowering the costs of engaging in particular scientific areas. Regardless of the channel, experience in an innovative process can alter both a scientists' overall productivity as well as their propensity to focus on a particular type of research insofar as they are able to, and have a preference to, select projects that will utilize their experience and benefit from the potential of increased productivity and lower costs in one domain over another.

In sum, scientists' ability to take advantage of specific learning opportunities may have important implications for their future research trajectories. An emerging body of prior work points to this idea, documenting the influence of formative experiences on scientists' careers and research output. For instance, Roche (2023) shows that doctoral students whose mentors are more engaged with start-ups and provide less intensive mentoring opportunities to students have lower research productivity after graduation. Azoulay et al. (2021) examine recently graduated medical doctors who received a laboratory research fellowship at the U.S. National Institutes of Health. The authors find that doctors who took part in the fellowship had significantly higher rates of engagement in research, in particular in translational research, later on in their careers compared to the applicants who passed screenings for the program but failed to find a laboratory match. Azoulay et al. (2017) show that scientists' attitudes to research commercialization can be imprinted in early career experiences with advisers. Further, Azoulay et al. (2009) show that scientists tend to produce more commercially oriented research following a patenting experience, driven in part by the development of new relationships with industry based researchers which exposes them to new questions.

Moreover, the effect of experience on scientists' future research trajectories may be limited by a range of contextual factors (Argote and Miron-Spektor, 2011). Factors such as individuals' networks, access to resources, and organizational structures may affect how they are able to leverage learning opportunities

(Bunderson and Boumgarden, 2010; Reagans and McEvily, 2003; Contu and Willmott, 2003). In the specific context of scientific research, scientists may be reliant on other actors to obtain external research funds and attract collaborators with complementary knowledge to complete a project (Wuchty et al., 2007; Haeussler and Sauermann, 2020; Jones, 2021; D'Este and Perkmann, 2011; Roach and Sauermann, 2010). In turn, this may affect scientists' ability to apply the knowledge and skills gained from one experience to subsequent projects. In environments with significant resource and institutional constraints, such as Sub-Saharan African science, these constraints may be more likely to bind than in environment where research funds are more munificent and there is a more extensive network of potential collaborators.

Not only might opportunities to apply accumulated knowledge and skills be limited, but the combination of the importance of learning through experience and scarcity of research resources leads to a tension for the researcher. On the one hand, if opportunities to develop expertise are rare, scientists may need to focus their research on areas where they have had the chance to enhance their knowledge and skills in order to maximize their research quality and impact. In settings where specialized skills are relatively rare, but crucial for a particular type of project, the returns to being a specialist may be high as a scientist becomes a crucial member of research teams (Teodoridis, 2018). On the other hand, by specializing narrowly into one area of research, scientists may miss opportunities to access knowledge and resources from other domains and be limited in the range of projects they can undertake (Conti et al., 2013; Nagle and Teodoridis, 2020).

We focus our hypotheses on a setting in which experience is likely to be important for scientists' future research careers, but where there may be limits to the benefits from experience: clinical trials in Sub-Saharan Africa.

2.2 The role of clinical trial experience on follow-on research amongst developing country scientists

Developing new and more effective treatments for high impact diseases affecting the Global South are widely seen as global grand challenge (George et al., 2016; Howard-Grenville, 2021; George et al., 2024). These are highly complex, long-term, global challenges that are difficult to overcome precisely because they require actors with disparate interests to work together to identify and mitigate barriers to progress (Couture et al., 2023; Ferraro et al., 2015). For example, Sustainable Development Goal 3 commits U.N. member states "[t]o ensure healthy lives and promote well-being for all at all ages." A core implementation target of this goal is to "[s]upport the research and development of vaccines and medicines for the communicable and noncommunicable diseases that primarily affect developing countries." Such a goal, by its very nature, requires significant coordination across the wide range of globally distributed actors involved in basic scientific research, early-stage drug development, clinical trials, and the manufacturing and distribution of pharmaceutical products. Universities and research funders, pharmaceutical companies, regulatory agencies, governments, and non-profit organizations all have important roles in the process of generating and translating scientific ideas into products that benefit patients globally.

Clinical trials are a critical part of the development process for new drugs, vaccines and diagnostics. Before a company or non-profit can access a market with a new drug or vaccine, they must go through a process of registration and regulation. As a part of this process, governments, or international regulatory agencies, require evidence that the drug is both safe and efficacious in humans. The evidence for this comes from clinical trials. There are various phases of clinical trials that every drug must go through. Phase I trials test for safety and tolerability of the drug in healthy volunteers, Phase II trials assess preliminary efficacy of the drug, while Phase III clinical trials assess the safety and efficacy in a larger population, and are often carried out at several sites, spanning the globe. Lastly, Phase IV trials typically take place after regulatory approval and aim to provide evidence on longer-term side effects.

Clinical trials from Phase II onwards are enormously expensive (DiMasi et al., 1991), costing up to half of all drug development costs. The median cost of a clinical trial is around USD $19 million (Moore et al., 2018). The enrollment of willing patients, specialized skills, staff and equipment are just some factors driving high costs of clinical research around the world.

2.2.1 Global patterns of clinical trials

The cost, time and difficulty of running a clinical trial varies significantly around the world (Qiao et al., 2019). The location of experienced staff, sensitized patient groups and specialized equipment and procedures, or trial capacity, is not evenly distributed, as evidenced in part by the observable variation in the extent to which countries host clinical trials. Recent evidence (Thiers et al., 2008) suggests that costs and time of the trial can vary widely from country to country, and firms seek to place their trials to minimize costs and time. Using data from ClinicalTrials.gov, we show that the proportion of trial sites around the world for all clinical trials is changing somewhat between 2000 and 2015 (Figure 1). That said, despite being home to more than 1 billion people, and home to much of the world's disease burden, particularly 'neglected tropical diseases' which account for a high proportion of disease and death in these regions, Sub-Saharan Africa hosts only a small fraction of global trials. Even within the region, there is variation in trial density (Figure 2). Beyond absolute volume of trials, the main partners and sponsors on trials vary around the world, with a large proportion of trials in Sub-Saharan Africa being sponsored by academic partners, as opposed to industry sponsors.[6]

[Figure 1 about here.]

[Figure 2 about here.]

This under-representation of trials has implications for the development of drugs and vaccines for problems affecting these locations. It is important to have trial capacity in disease-endemic settings in order to complete product development for these diseases.

[6]https://www.clinicaltrialsarena.com/features/academic-commercial-clinical-trials/?cf-view

2.2.2 Opportunities to gain experience in clinical trials

Most of the funding for clinical trials in developing countries comes from external organizations, including foreign governments, non-profit organizations, and pharmaceutical companies, and most clinical trial projects involve foreign collaborators. Funders and collaborators require that clinical trials are executed successfully and need to work with researchers who have the necessary scientific knowledge, research skills, and project management skills on the ground. Such necessary skills include sample collection, data entry, laboratory analysis, trial management, patient recruitment, and regulatory approval navigation (Franzen et al., 2017; Alemayehu et al., 2018). As an interviewee noted to one of the authors: *"one of the major reasons why industry doesn't want to site trials in developing countries. There aren't investigators on whom they can rely…."*

One plausible way for Sub-Saharan African scientists to acquire this expertise would be from experience, for instance, by working on a trial, particularly alongside scientists who have experience running clinical trials. Science is a context in which tacit knowledge gained learning from others is highly important for technical research skill development and for understanding how to design and manage research projects (Polanyi, 1958; Senker, 1995; Ravetz, 1971). For example, through working on a research project, scientists learn how to take and manage samples, carry out laboratory analysis, record results, and prepare data, and over time may learn other managerial skills such as coordinating others' workstreams and managing relationships with external stakeholders. In addition, a project experience can lead to improvements in infrastructure, development of relevant networks, increased visibility and elevated awareness of future relevant opportunities in the space. In turn, providing scientists with an opportunity to gain an experience may have a significant impact on their subsequent research decisions. For example, Franzen et al. (2017) cite a Cameroonian scientist who explained to the authors that: *"Getting exposed to different aspects of research and working with different groups of people is an experience you really can only have if you are part of it [clinical trial]. Your knowledge increases, your understanding, you have to think deeper. Interacting with high profile professors who are very experienced, I learned a lot."*

Opportunities to learn from experience may not only improve scientists' task-associated capabilities, but also help scientists develop knowledge about how to apply those knowledge and skills in different contexts (Chase and Simon, 1973; Ericsson et al., 1993; Simon, 1991; Dane, 2010; Greenwood et al., 2019). Knowledge for running clinical trials may extend to other types of research study that involves similar tasks, including sample collection, data entry, laboratory analysis, trial management, and patient recruitment.

We hypothesize that participating in clinical trials provides opportunities for scientists to develop relevant capabilities related to managing clinical trials and that can be applied to related types of research. Scientists will benefit by increasing their research focus on areas where these capabilities are most relevant, rendering them more likely to pursue future clinical trials research opportunities. In addition, in our context there may still be relatively few opportunities to work on clinical trials. As a result, following trial experience, we expect scientists will also lead pursue research projects that applies their capabilities to proximate

types of research, such as clinically focused projects on a disease in which they gained experience. This leads to Hypothesis 1:

H1. Developing country scientists with clinical trial experience are more likely to pursue subsequent clinical trials and more applied, disease focused research.

2.3 Who benefits from clinical trial experience?

Much of scientific knowledge is cumulative; scientists need to master basic ideas and skills in order to develop advanced research capabilities, as exhibited by the extensive time invested in doctoral and post-doctoral training (Jones, 2009). In the context of clinical trials, if expertise is slow to develop and requires extensive experience, this may limit the impact of clinical trial experience on future research direction. In this case, participation in one trial may not give scientists the requisite knowledge and skills, or the opportunity to establish more developed working relationships, to be an attractive partner for carrying out future trials. In these instances, there would be cumulative advantages to those with more prior clinical trials experience (Merton, 1968; Bol et al., 2018).

On the other hand, there could be diminishing marginal returns to clinical trials experience. In some instances, for example, if the tasks required are relatively more routine or the infrastructure and networks are straightforward to develop, scientists may not need extensive clinical trial experience to have sufficient capabilities to shift their research towards more clinical trials and related projects. Instead, the necessary capabilities may be gained relatively quickly as scientists gain experience participating in trials. In our research, one clinical trial participant noted: *"I was site investigator for multiple sites... We collected data during the process and it handed over to the NIH team. I think it was sent back to America... we realized that it wasn't that difficult to draw up a research program and execute it."* For scientists who already have complementary scientific knowledge, the pace of learning additional skills for clinical trials tasks may be relatively rapid. Franzen et al. (2017) cite the comments of a head of a scientific department that participated in a clinical trial consortium in Cameroon who noted: *"Participating [in X consortium] has given us this opportunity to build collaborations with very good researchers. People now know that we exist, and that is good. We have the capacity now to go and develop. All my students are going to learn clinical training. I don't have any problem with that now I have an infrastructure."*

If learning is rapid, the gains to additional trial experience may be relatively small. Even if there are meaningful gains to additional experience for those scientists with prior clinical trial experience, they may have other binding constraints that limit their ability to apply the knowledge and skills gained in subsequent trials. Those who have made a prior choice to pursue clinical trials research may have individual-level capacity constraints in the number of trials they can—or would choose to—carry out at a point in time. Running clinical trials may be a time-consuming type of research with a ceiling on the number of trials can run. Alternatively, in an environment where research funding is uncertain, scientists may choose not to specialize very narrowly into one type of research. More widely applicable human capital—i.e., from hav-

ing more of a generalist research program—may be valuable in these settings because it enables a scientist to take advantage of a wider range of opportunities in their environment (Nagle and Teodoridis, 2020). As a result, there may be limits on the extent to which scientists choose to focus efforts on clinical trials at the expense of other types of research. The discussion in this section leads to our second hypothesis:

H2. Prior experience in clinical trials negatively moderates the relationship between clinical trial experience and subsequent clinical trials.

3 Setting, Empirical Strategy and Data

3.1 The European and Developing Countries Clinical Trial Partnership

In a recognition of a need to run trials and develop trial capacity in Sub-Saharan Africa, The European and Developing Countries Clinical Trial Partnership (EDCTP) was established in 2003. EDCTP is an EU funded partnership between 15 European countries and 25 African countries. Their mission is to support collaborative research and accelerate the development of new or improved medical interventions for the identification, treatment and prevention of infectious diseases in Sub-Saharan Africa, through all phases of clinical trials, with an emphasis on Phase II and III trials.

During the first program (EDCTP-1, 2003-2015), EDCTP supported clinical research and career development, PhD and masters fellowships on treatment drugs, vaccines and diagnostics, focusing on HIV/AIDS, malaria and tuberculosis. With a budget of over 1 billion euros, EDCTP-1 was the first, and largest program to support clinical trials in Sub-Saharan Africa at the time. Prior to the EDCTP-1 program there were limited opportunities to participate in clinical trials in Sub-Saharan Africa.

EDCTP-1 funded various types of project, including providing full funding for collaborative clinical trials, fellowships, MSc and PhD scholarships, and a handful of ethics, regulatory and network strengthening projects. We focus on the funded projects that can be identified as a single trial: either collaborative trials themselves, or fellowship/scholarship projects that provide funding to support either a senior, or an early career individual within an identifiable clinical trial.

One such sponsored trial was a trial known as TaMoVac-01, which was a Phase I/II randomized controlled trial in adults in Tanzania to study the safety of a HIV vaccine candidate. Started in 2008, this 4 year project was led by Muhammad Bakari at the Muhimbili College of Health Sciences in Tanzania, with collaborators from Europe and other African institutions.

3.2 Empirical Strategy

In this paper we estimate the relationship between being involved in an EDCTP sponsored clinical trial and subsequent trial involvement and research agenda of African scientists. Our empirical strategy compares the change in outcomes of scientists who participate in a trial to that of matched control scientists

in a difference in differences framework.

3.3 Sample construction

3.3.1 EDCTP trial participants - treated scientists

We identified 1,198 scientists who participated in an EDCTP trial between 2005 and 2014, were affiliated with an African institution, and could be matched to a publication record in the Elsevier Scopus publication database. Out of the 1,190 scientists, 433 participated in HIV trials, 356 in malaria trials, and 434 in TB trials (with some overlap between TB and HIV trials in instances where trials were multi-disease focused). As for location of the researchers, 22 percent were from South Africa, 10 percent from Tanzania, 11 percent from Uganda and 9 percent from Kenya. Overall, 880 of the 1,198 scientists had a publication record prior to the trial, and are included in our analysis.

3.3.2 Control scientists

To account for trends over time and over a scientist's career, we incorporate a control group of scientists who are also affiliated with African institutions, but who are not involved in an EDCTP trial. We use the Elsevier Scopus publication database, and institutional affiliations in publication records, to extract the full set of publishing scientists in Africa between 2005 and 2015. From this set of scientists we extract a smaller sample who are carefully matched with the treated scientists. Namely, for each treated scientist we use a coarsened exact matching procedure to identify a set of control scientists who are precisely matched on variables such as career age, OECD collaborations, disease focus of research, applied nature of research, clinical trial involvement and institutional level size and trial involvement in the three years prior to the treated scientist's EDCTP trial. Unmatched treated scientists are discarded. The matching procedure leaves us with 794 treated scientists (or just over 90 percent of our original EDCTP participant sample), and 144,475 control scientists. Each control scientist is assigned a counterfactual EDCTP trial which corresponds to the trial participated in by their closest matched treated peer.

3.4 Variables and Measures

The data in our sample is constructed from three main sources. We draw on clinical trials data from ClinicalTrials.gov, publications data from Scopus, and link publications in Scopus to the keywords indexed to papers in the PubMed database. The ClinicalTrials.gov database is managed by the US National Library of Medicine within the National Institutes of Health and was created in 1997 following the Food and Drug Administration Modernization Act.[7] There are currently more than 470,000 ongoing and completed clinical trials in the database, each with a unique trial identifier. We use the ClinicalTrials.gov database to identify

[7] After additional requirements were introduced in 2007, the Sponsor or Principal Investigator of a clinical trial has been required by law to register and report results for clinical trials of drugs, biologics, and devices that are subject to FDA regulation with the exception of certain Phase 1 trials.

clinical trials involving the treated and control scientists in our sample using a matching procedure that included investigator name and publication outputs from trials that we linked to Scopus records. We also identify the location of each clinical trial site in the database using a key word search, and use the results of this search to generate measures of annual clinical trials taking place at the country level.

We augment this data by building a rich bibliographic dataset based on scientists' publications. We use the Elsevier Scopus author identifier to extract each sample scientist's publications recorded in this database. We use the bibliographic data in Elsevier Scopus to identify the diseases of focus in scientists' papers, based on keywords in their abstracts, the publication outlet of each paper, scientists' institutional affiliations over their careers, and the affiliations of their co-authors. We also use the PubMed identifiers of papers recorded in Elsevier Scopus to link these publications to MeSH terms, or descriptor terms used to organize concepts in research.[8] Each paper in our sample is placed at a particular point in the space of scientific concepts based on its content. This allows us to track changes in the direction of scientists' research over time according to changes in the location of their papers in MeSH space.

We combine the data from each source to create a panel dataset at the scientist-year level. This contains scientists' career histories with yearly observations of variables that measure scientists' participation in clinical trials, co-authorship relationships, and publication outcomes for the four years before, and ten years after the EDCTP trial (or counterfactual trial in the case of the control scientists).

3.4.1 Measures

Our first set of dependent variables build on the data from ClinicalTrials.gov. We are interested in how participation in an EDCTP trial is associated with future clinical trial participation. Our primary dependent variable to test this relationship is the number of clinical trials in which scientist i participates in year t (based on the year of a trial's start date). Second, we create separate dependent variables according to whether the drugs or vaccines in a clinical trial are targeting a disease that was a focus of the initial EDCTP program that the focal scientist was involved in, or a different disease. This enables us to examine how experience in an EDCTP-supported trial is linked to scientists participating in clinical trials for other types of disease. We also restrict the dependent variable to trials with sponsoring organizations from different regions to analyze how EDCTP participation affects the integration of African scientists into the global system of clinical trials.

Second, we create a set of dependent variables to track changes in scientists' publications over time. We are interested in the changes in the rate and direction of scientists' after participating in the EDCTP program and any changes in their patterns of collaboration with other scientists. We create a series of variables to measure these using bibliographic data from Scopus. To measure changes in the rate of scientists' research,

[8]The Medical Subject Headings vocabulary is managed by subject-specific experts at the National Library of Medicine. There are approximately 30,000 descriptor terms in the MeSH vocabulary, which are used to organize concepts in medicine and life sciences research into a hierarchical tree format. Indexing is independent of article authors. Terms are assigned by indexers at the NLM who select them based on a specific protocol. The OpenAlex database records the MeSH terms of included publications, which we match to our Scopus sample using the PubMed IDs common to both databases.

we create variables measuring: the number of publications in Scopus on which scientist i is an author in year t; the number of publications on which scientist i is an author in year t weighted by journal impact factor; and the number of publications on which scientist i is the first or last author in year t. To measure changes in the direction of scientists' research we create variables that separately count scientist i's number publications in year t according to the diseases that are the focus of each publication. We also examine changes in the commercial relevance of scientists' research. We define an applied scientific publication as a publication in a journal for which the Journal Commercial Impact Factor is strictly positive (Bikard and Marx, 2020), basic otherwise. To measure changes in the collaboration patterns of scientists we examine the location of their co-authors. In particular, we create two separate variables counting the number of scientist i's publications in year t that involve co-authors from OECD and non-OECD countries respectively. Again, we apply the inverse hyperbolic since transformation to the raw number of each type of trial when constructing the dependent variables.

Finally, we create dependent variables based on the MeSH terms that are independently indexed to scientist i's publications by the National Library of Medicine. We use these to analyze the extent to which scientists build new knowledge and skills through trial participation. First, we restrict our sample to the set of publications that are identified as being linked to a specific EDCTP supported trial. We create a variable that measures the share of MeSH terms indexed to one of these trial-linked publications by scientist i, which had not been indexed to any of scientist i's publications from prior years. This provides a proxy measure of how far scientist i is developing new knowledge and skills in the EDCTP clinical trial. Second, we create a dependent variable that measures the number of MeSH terms that were used in post-treatment publications more generally that were a. used for the first time in EDCTP trials, and b. used in an EDCTP trial, but not for the first time for that scientist.

3.4.2 Descriptive statistics

Table 1 presents any differences in clinical trial participation, publications, institutions, co-authorship relationships, and experience between treated and control scientists in the three years prior to the EDCTP trial in a three-year period prior to the year in which the treated scientists in a given stratum first participated in the EDCTP program (or counterfactual).

There is no significant difference between treated and control scientists in their individual rate of prior clinical trial participation, the number of clinical trials in which scientists at their organization participated, the size of the institution with which they are affiliated (based on the number of researchers working at that organization), their number of co-authors from OECD countries, and their career age at the time of the EDCTP trial. However, treated scientists have a slightly higher number of publications in the three years prior to the treatment event. Treated scientists also have more publications than the control scientists on HIV, malaria, and tuberculosis, which were the three diseases of focus in the EDCTP grant program. In online appendices we run a number of tests to confirm that this difference is not driving the observed

effects. Moreover, the significant differences in publications numbers between treated and control scientists are only present among scientists matched in the tuberculosis arm of the EDCTP program. Our results are robust to excluding all tuberculosis trial scientists and their matched controls from the analysis.

[Table 1 about here.]

4 Results

We now turn to examining how scientists' participation in an EDCTP funded trial corresponds to future participation in clinical trials and any changes in their scientific research output. We use difference-in-differences analysis to analyze treated scientists' outcomes after participating in the program to those of matched control scientists.

Our strategy allows us to account for individual level heterogeneity, career age trends and year trends through the use of fixed effects. Specifically, our core regression model applied to a scientist year level dataset includes scientist and year fixed effects to control for time-invariant heterogeneity linked to scientists' characteristics and wider variation in trial or publication outcomes over time. We also include a 'career age' fixed effect, defined as the number of years since a scientist's first publication recorded in the Elsevier Scopus database to control for changes in research over career lifecycles. We also include a fixed effect that measures the number of years before or after an EDCTP grant is received for each treated scientist and their matched controls. Our key identifying assumption is that the treated and control scientists would follow the same trajectory in terms of clinical trial involvement and publication outcomes in the absence of the treatment.

Formally, we estimate the following regression model for each of our dependent variables:

$$y_{it} = EDCTPGrantee \times PostGrant_{it} + s_i + t_t + c_{it} + d_{it} + \varepsilon_{it} \tag{1}$$

The variable $EDCTPGrantee \times PostGrant$ is our treatment variable which takes the value of 1 if an individual is involved in an EDCTP trial and the observation year is post EDCTP trial. Finally, s_i represents the scientist fixed effect, t_t represents the year fixed effect, c_{it} represents the career age fixed effect, and d_{it} represents the time to/from treatment fixed effect for the scientists. Standard errors are clustered at the scientist level.

4.1 Participation in Clinical Trials

Our first set of results concerns the relationship between participating in the EDCTP program and scientists' future participation in clinical trials. Our primary dependent variable is the number of clinical trials in which scientist i participates in year t.

[Table 2 about here.]

The results in Column 1 of Table 2 show that there is a significant increase in the number of clinical trials per year in which scientists participate after EDCTP treatment. If we interpret the point estimate of the coefficient relative to the sample mean, scientists who participated in the EDCTP program are about three times as likely as matched controls to participate in a clinical trial in a given year. In Column 2, we show that this does not seem to be driven by changes in the overall research productivity of scientists. If we control for the number of publications a scientist has in a given year, the estimates are very similar. In online appendices we illustrate the robustness of these results to alternative functional forms, including Poisson maximum likelihood and linear probability models.

In an illustration of the dynamics of the effect, Figure 3 shows in the raw data how treated and control scientists have very similar rates of participation in clinical trials (excluding those supported by an EDCTP grant) before a scientist participates in the EDCTP program. However, following EDCTP participation, scientists begin to participate in other clinical trials at a higher rate than control scientists. Figure 4 show how the magnitude of the difference in trial participation changes over time after scientists participate in the EDCTP program using an event study. There is not an immediate divergence in trial participation. The difference between treated and control scientists' participation in trials grows over time. This is not surprising. Clinical trials take multiple years to design, recruit, and run and so we would expect a lag to allow for the EDCTP trial to finish and subsequent trials to be initiated.

[Figure 3 about here.]

[Figure 4 about here.]

We next examine how participation in the EDCTP program corresponds to participation in subsequent clinical trials across different diseases. If treated scientists are subsequently more likely to work on clinical trials in disease areas distinct from that of their initial trial, this would suggest that the knowledge and skills developed by participating in an EDCTP trial can be transferred to projects in other disease areas. The results in Columns 3 and 4 of Table 2 show that this is the case. Treated scientists are relatively more likely to participate in subsequent clinical trials for both the same disease as their initial trial, and other diseases.

Finally, in Columns 5 and 6, we show that increases in trial participation is reflected in the impact of scientists' trial-related research output. Treated scientists experience a significant increase in the production of trial publications (weighted by source normalized impact per paper, a quality measure). This suggests that treated scientists' increase in trial participation is not driven by low quality trials. In addition, we show that there is a small, and weakly significant change in the rate at which treated scientists are the first or last author on trial publications, implying that they are slightly more likely to occupy leadership positions on subsequent clinical trials.

In the event that there are positive spillovers to other scientists in the same institution, but who are not involved in the EDCTP trial, we run the same analysis excluding those in the same institution. The results provided in the online appendix show that our main estimations are a lower bound estimate, but that results remain qualitatively similar.

[Table 3 about here.]

The EDCTP program was funded by the European Union and involved collaborations between African scientists and European partners. If participating scientists develop knowledge and skills that transfer to projects funded by other organizations, we should see some evidence that scientists are more likely to work on clinical trials with new partners after participating in the EDCTP program. In Table 3, we examine changes in the rate at which treated scientists participate in clinical trials with sponsors from the EU, US, or Africa. There are significant increases in treated scientists' participation in clinical trials sponsored by EU and US organizations. The increase in participation in US-sponsored trials indicates that treated scientists are working on trials with new funders and suggests that the increase in scientists' subsequent trial participation is not driven by partner-specific relationships. However, we find little conclusive evidence of changes in scientists' rate of participation in Africa-sponsored clinical trials after the EDCTP program.

4.2 Changes in Research Direction

We now examine other changes in scientists' research after participating in the EDCTP program. For example, they may focus subsequent research more on applied topics or research on the same diseases if they acquired relevant knowledge and skills during the EDCTP program.

[Table 4 about here.]

The results in Columns 1 and 2 of Table 4 show that there is an increase in the rate at which scientists' publications have specific diseases as a topic after treatment, alongside a decrease in non-disease focused publications, even after controlling for overall publications. The results in Columns 3 and 4 show that this is driven entirely by treated scientists increasing the rate at which they publish research on their EDCTP trial diseases, with no change on publications focused on other diseases. Finally, the results in Columns 5 and 6 show that treated scientists experience a relative increase in the rate at which scientists' publications contain applied research and appears in more commercially-relevant journals and a decrease in basic research publications after participating in the EDCTP program.

4.3 Heterogeneity by Prior Trial Participation

Participation in a first clinical trial might be especially important to scientists if there are steep returns to an initial exposure to an experience, either because prior to an exposure, access to similar opportunities was

not possible or because there are diminishing returns to learning-by-doing in the context of clinical trials in Sub-Saharan Africa. If this were the case, we would expect to find that scientists who had previously participated in a clinical trial prior to participating in the EDCTP program would have a smaller increase in subsequent trials than scientists for whom the EDCTP program offered a first clinical trial experience. We augment our core model with interaction effects to capture prior trial experience. In particular, we create a triple interaction between the *Post* and *Grantee* variables with an indicator variable *PriorTrial*. This additional variable denotes whether scientist i had participated in a clinical trial recorded in the ClinicalTrials.gov database in the three years prior to EDCTP grant (or counterfactual). To saturate the model we also interact our *Post* variable with the *PriorTrial* variable. Since both the *Grantee* and *PriorTrial* variables are time-invariant, they are absorbed by the scientist fixed effects in the model. Formally, we estimate the model:

$$\begin{aligned} y_{it} = EDCTPGrantee \times PostGrant \times PriorTrial_{it} + \\ EDCTPGrantee \times PostGrant_{it} + PostGrant \times PriorTrial_{it} \\ + s_i + t_t + c_{it} + d_{it} + \varepsilon_{it} \end{aligned} \tag{2}$$

The results in Tables 6 and 5 replicate those from Table 2 using our triple interaction model. There is a clear pattern in which the relationship between EDCTP trial involvement and subsequent trial participation are smaller for scientists who participated in the EDCTP program and had prior clinical trial experience relative to those for whom the EDCTP program offered a first experience of clinical trials, and for those with fewer years of research experience more generally. Notably, the estimates in Column 1 of Table 6 suggest that scientists with prior trial experience had no observable increase in subsequent trial participation after the EDCTP grant. However, for scientists without prior trial experience, participation in the EDCTP program is associated with an increase in the rate of subsequent trial participation that is approximately 40% greater than that in the full sample estimates in Table 3.

[Table 5 about here.]

[Table 6 about here.]

In Table 7 we explore whether the heterogeneous results for those with and without prior trials is driven by career age retirement concerns, or by differences in overall productivity, exposure to research, research focus or prior networks of the focal scientist. We control for any differences in changes in trial participation by these additional scientist level factors, and whilst not exhaustive, this provides limited evidence of alternative explanations driving the results.

[Table 7 about here.]

4.4 Evidence of Learning-by-doing

We now turn to providing more evidence of the mechanisms through which this individual-level heterogeneous effect is taking place. On the one hand, scientists may develop relevant knowledge and skills by participating in a first clinical trial in the EDCTP program, which facilitates participation in subsequent trials with other sponsors. In this case, learning-by-doing would build individual scientists' capabilities for clinical trial research. The knowledge and skills scientists developed by first trial participation would then be transferable to future projects with other partners. On the other hand, participating in their first trial may have helped scientists increase their network of collaborators or gain greater global visibility, which would help access partners with whom they could carry out future trials. In this case, trial participation would not be leading to individual-level learning, but rather provide African scientists with greater access to the networks of global scientific research. In turn, this would help them build partnerships for future trials.

To test these competing mechanisms, we carry out two sets of additional analyses. First, we examine whether participation in a trial through the EDCTP program is associated with increased scientific visibility (in terms of scientists' raw and impact-factor adjusted publications) or with increased numbers of collaborations with global scientists. Second, we analyze the MeSH terms indexed to the specific publications of our treated scientists to examine whether there is evidence that scientists develop knowledge of new concepts or skills by participating in a clinical trial that is subsequently re-used in their future research.

[Table 8 about here.]

Since we only find an increase in the rate of future trial participation among scientists working on a first clinical trial, if this increase is driven by a visibility or network mechanism, we should find evidence that these scientists also have larger increases in their research productivity or their number of collaborations relative to those with prior trial experience. The results in Table 8 show that this is not the case. The results in Columns 1 and 2 indicate that scientists with and without prior trial experience have similar increases in the annual number of publications. We also find no evidence that first-time trial participants extend their network of collaborators more than repeat trial participants. Interestingly, the results in Column 3 do show that repeat trial participants increase the rate at which they are the first or last author on publications after EDCTP participation by more than first time trial participants. This indicates that these scientists may be building on the EDCTP grant to take on more leadership roles in research projects, rather than increasing their rate of trial participation. The results in Columns 4 and 5 show that there is also no evidence of significant differences in the rate at which first time and repeat trial participants collaborate with other scientists.

Next, we examine the concepts associated with scientists' publications to examine whether there is evidence of learning-by-doing. We use the MeSH terms indexed to the treated scientists post-treatment papers to analyze whether scientists' research involves new-to-the-scientist concepts in their publications

linked to EDCTP trials (i.e., terms that had not been indexed to scientist *i*'s research in prior years). If first participation in a clinical trial is associated with learning-by-doing, we should find first time trialists have a greater share of new-to-the-scientist terms linked to their trial publications compared to repeat trialists.

[Table 9 about here.]

Our unit of analysis is the publication, and in Table 9 Column 1 our sample is just publications linked to treated scientists' EDCTP trial. In Columns 2 to 5, the sample is all post-treatment publications of treated scientists. The results in Column 1 of Table 9 show that treated scientists with prior trial experience have a lower share of brand new (to them) MeSH terms indexed to their EDCTP trial publications as compared to treated scientists without prior trial experience. That is, their trial publications cover new scientific concepts at a higher rate than scientists participating in an EDCTP trial who have prior trial experience. Those with prior trial experience appear to re-use existing knowledge and skills at a higher rate.

In Column 3, we show that, conditional on the number of new-to-the-scientist MeSH terms indexed to their trial publications, first-time trialists' subsequent research is indexed to these new-to-the-scientist concepts at a higher rate than those with prior trial experience. This suggest that novice trialists' research direction shifts more to incorporate new knowledge or skills developed during a trial than the research direction of scientists with prior experience of clincial trials. Conversely, in Column 4 and 5, we show that the rate at which repeat trialists' subsequent research is indexed to MeSH terms that were associated with both their EDCTP trial publications and their pre-EDCTP publications is higher than for first time trialists. This is consistent with repeat trialists having a greater range of relevant knowledge and skills before participation in the EDCTP program that is relevant to their ongoing research. The MeSH terms corresponding to these concepts would then be indexed to the subsequent publications.

4.5 Heterogeneity by Country Trial Capacity

Lastly, we explore heterogeneity by focal scientist's country trial capacity to assess the extent to which institutional constraints might be limiting benefits from an experience. Unfortunately, data on the institutions required to efficiently run trials in countries (for example, regulatory and ethical bodies) is not available. Thus we use a proxy for these institutions which is the number of trials taking place in a focal scientist's country prior to the EDCTP trial. Namely, in Table 10 we interact the post grant and EDCTP grantee dummy by an indicator that takes the value of 1 if a focal scientist is affiliated with an African country that has above median number of trials in the three years prior to the EDCTP trial in the same calendar year, 0 otherwise.

We find that on average the number of trials is influenced by country level capacity, and in particular, more senior scientists in countries with more prior trials experience some increase in trials. This is in contrast to scientists in countries with lower capacity, whereby only scientists with no prior experience increase their trial participation. This implies a ceiling to trial participation in some environments. In

online appendices we show that this result is robust to including variation according to country level GDP, and to variation in terms of individual scientist's productivity and research agenda. Country level trial capacity also doesn't appear to influence overall changes in research productivity or networks. Together, this provides suggestive evidence that country level trial capacity can play a role in specialization patterns following an domain specific experience.

[Table 10 about here.]

5 Discussion

In this paper, we examine how a scientist's experience gained through participating as a researcher in a clinical trial corresponds to follow on involvement in clinical trials and their broader research trajectory. We examine the EU's EDCTP grant program to provide evidence that participating in a clinical trial is strongly associated with higher likelihoods of future trial participation. We find that this increase is largest for scientists participating in a clinical trial for a first time. We also provide indicative evidence of learning-by-doing. Where scientists have participated in other trials prior to treatment, their publications linked to a focal clinical trial involve a relatively lower share of novel scientific concepts than novice's trial-linked publications. Finally, we document that subsequent clinical trial involvement is greater for scientists based in countries with high levels of prior trials.

Our results contribute to the literature on the determinants of research direction. While prior research on the determinants of research direction has explored the role of funding availability (Myers, 2020), exposure to problems (Truffa and Wong, 2022; Fry, 2023), peers (Catalini, 2018; Azoulay et al., 2019), and access to data and tools (Nagaraj et al., 2020; Furman and Teodoridis, 2020), in this paper we highlight the role of experience in shaping the research trajectory of scientists. Our key contribution is to propose that in some contexts, path dependency of science is not a given. Namely, we show that environmental constraints limit the extent to which experience shapes research direction. We also contribute to the literature on the role of formative experiences on scientists' research trajectories (Roche, 2023; Shibayama, 2019; Azoulay et al., 2021; Conti et al., 2013; Azoulay et al., 2017). In line with this literature, we show that, in some contexts, a formative experience can shape the careers of scientists, specifically by shaping their trajectory of research. This work also contribute to research demonstrating the important role of institutional environments in innovative outcomes (Fry and Furman, 2023; Vasudeva et al., 2013; Wang, 2015) by hypothesizing and documenting the importance of national innovative capacity in determining scientists' ability to apply the benefits gained from research experiences.

More widely, we respond to calls for researchers in the strategy and management fields to contribute insights from our disciplines to help find solutions to global grand challenges, and, in particular, to those in global health (George et al., 2016; Howard-Grenville, 2021; George et al., 2024; Arslan and Tarakci, 2022). We shed light on how interactions between the human capital of the people working to solve grand challenges

and the institutional environment shape individuals' ability to apply their human capital to solving grand challenges. In the context of global clinical trials, we show that coordinated action at both the micro- and macro-level will be necessary to increase clinical trial capacity in resource constrained environments. On the one hand, there is a need for coordination across organizations funding scientific research and drug development to help scientists learn and fill skills gaps so that there are people with the individual-level capabilities to carry out clinical trials. On the other hand, these individual-level capabilities will only be applied to trials that seek to solve global health challenges if governments in resource constrained countries are able to create supportive infrastructure to attract partners from other countries to work with domestic scientists on clinical trials.

Our results have some important limitations. First, our results can not be interpreted as causal. However, given the nature of the under-studied context in this study, and the detailed evidence we provide on heterogeneity and mechanisms driving any observed effects, we maintain that these findings provide an important step forward in our understanding of the drivers of scientists research trajectories. That said, future research should seek to exploit additional experimental and quasi-experimental methods to further our understanding of this highly understudied population of scientists. Second, we cannot infer the optimal trajectory of research for sample scientists. Future work should seek to understand the broader effects of experience and the follow-on research trajectories of scientists on macro-level labor force trends, health and economic outcomes.

Overall, the findings in this paper suggest that efforts to improve clinical trials capabilities to solve health challenges affecting the Global South will require both coordinated investments in individuals to provide the necessary experience to work on trials and efforts to improve the regulatory and policy environments to attract trial resources to scientists in countries with resource and institutional constraints.

References

Adepoju, P. (2022). Africa's future depends on government-funded r&d. investing in research is key pathway to development on the continent. *Nature Africa*.

Adler, P. S. and Clark, K. B. (1991). Behind the learning curve: A sketch of the learning process. *Management science*, 37(3):267–281.

Agarwal, R. and Ohyama, A. (2013). Industry or academia, basic or applied? career choices and earnings trajectories of scientists. *Management Science*, 59(4):950–970.

Agrawal, A., Goldfarb, A., and Teodoridis, F. (2016). Understanding the changing structure of scientific inquiry. *American Economic Journal: Applied Economics*, 8(1):100–128.

Alemayehu, C., Mitchell, G., and Nikles, J. (2018). Barriers for conducting clinical trials in developing countries-a systematic review. *International journal for equity in health*, 17:1–11.

Argote, L. (2012). *Organizational learning: Creating, retaining and transferring knowledge*. Springer Science & Business Media.

Argote, L., Beckman, S. L., and Epple, D. (1990). The persistence and transfer of learning in industrial settings. *Management science*, 36(2):140–154.

Argote, L., Lee, S., and Park, J. (2021). Organizational learning processes and outcomes: Major findings and future research directions. *Management Science*, 67(9):5399–5429.

Argote, L. and Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization science*, 22(5):1123–1137.

Arslan, B. and Tarakci, M. (2022). Negative spillovers across partnerships for responsible innovation: Evidence from the 2014 ebola outbreak. *Journal of Management Studies*, 59(1):126–162.

Azoulay, P., Ding, W., and Stuart, T. (2009). The impact of academic patenting on the rate, quality and direction of (public) research output. *The Journal of Industrial Economics*, 57(4):637–676.

Azoulay, P., Fons-Rosen, C., and Zivin, J. S. G. (2019). Does science advance one funeral at a time? *American Economic Review*, 109(8):2889–2920.

Azoulay, P., Greenblatt, W. H., and Heggeness, M. L. (2021). Long-term effects from early exposure to research: Evidence from the nih "yellow berets". *Research Policy*, 50(9):104332.

Azoulay, P., Liu, C. C., and Stuart, T. E. (2017). Social influence given (partially) deliberate matching: Career imprints in the creation of academic entrepreneurs. *American Journal of Sociology*, 122(4):1223–1271.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49.

Bikard, M. and Marx, M. (2020). Bridging academia and industry: How geographic hubs connect university science and corporate technology. *Management Science*, 66(8):3425–3443.

Bol, T., De Vaan, M., and van de Rijt, A. (2018). The matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19):4887–4890.

Bunderson, J. S. and Boumgarden, P. (2010). Structure and learning in self-managed teams: Why "bureaucratic" teams can be better learners. *Organization Science*, 21(3):609–624.

Catalini, C. (2018). Microgeography and the direction of inventive activity. *Management Science*, 64(9):4348–4364.

Chan, T. Y., Li, J., and Pierce, L. (2014). Learning from peers: Knowledge transfer and sales force productivity growth. *Marketing Science*, 33(4):463–484.

Chase, W. G. and Simon, H. A. (1973). Perception in chess. *Cognitive psychology*, 4(1):55–81.

Conti, A., Denas, O., and Visentin, F. (2013). Knowledge specialization in ph. d. student groups. *IEEE Transactions on Engineering Management*, 61(1):52–67.

Contu, A. and Willmott, H. (2003). Re-embedding situatedness: The importance of power relations in learning theory. *Organization science*, 14(3):283–296.

Couture, F., Jarzabkowski, P., and Lê, J. K. (2023). Triggers, traps, and disconnect: How governance obstacles hinder progress on grand challenges. *Academy of Management Journal*, 66(6):1651–1680.

Dane, E. (2010). Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. *Academy of Management Review*, 35(4):579–603.

Darr, E. D., Argote, L., and Epple, D. (1995). The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises. *Management Science*, 41(11):1750–1762.

Dentoni, D., Bitzer, V., and Schouten, G. (2018). Harnessing wicked problems in multi-stakeholder partnerships. *Journal of Business Ethics*, 150:333–356.

DiMasi, J. A., Hansen, R. W., Grabowski, H. G., and Lasagna, L. (1991). Cost of innovation in the pharmaceutical industry. *Journal of health economics*, 10(2):107–142.

D'Este, P. and Perkmann, M. (2011). Why do academics engage with industry? the entrepreneurial university and individual motivations. *The Journal of Technology Transfer*, 36:316–339.

Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.

Epple, D., Argote, L., and Devadas, R. (1991). Organizational learning curves: A method for investigating intra-plant transfer of knowledge acquired through learning by doing. *Organization science*, 2(1):58–70.

Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3):363.

Ferraro, F., Etzion, D., and Gehman, J. (2015). Tackling grand challenges pragmatically: Robust action revisited. *Organization studies*, 36(3):363–390.

Fong Boh, W., Slaughter, S. A., and Espinosa, J. A. (2007). Learning from experience in software development: A multilevel analysis. *Management science*, 53(8):1315–1331.

Franzen, S. R., Chandler, C., and Lang, T. (2017). Health research capacity development in low and middle income countries: reality or rhetoric? a systematic meta-narrative review of the qualitative literature. *BMJ open*, 7(1).

Franzoni, C., Scellato, G., and Stephan, P. (2011). Changing incentives to publish. *Science*, 333(6043):702–703.

Fry, C. and Furman, J. L. (2023). Migration and global network formation: Evidence from female scientists in developing countries. *Organization Science*, (Articles in Advance).

Fry, C. V. (2023). Crisis and the trajectory of science: Evidence from the 2014 ebola outbreak. *Review of Economics and Statistics*, 105(4):1028–1038.

Fry, C. V., Lynham, J., and Tran, S. (2023). Ranking researchers: Evidence from indonesia. *Research Policy*, 52(5):104753.

Furman, J. L. and Teodoridis, F. (2020). Automation, research technology, and researchers' trajectories: Evidence from computer science and electrical engineering. *Organization Science*, 31(2):330–354.

Gambardella, A., Ganco, M., and Honoré, F. (2015). Using what you know: Patented knowledge in incumbent firms and employee entrepreneurship. *Organization Science*, 26(2):456–474.

George, G., Fewer, T. J., Lazzarini, S., McGahan, A. M., and Puranam, P. (2024). Partnering for grand challenges: A review of organizational design considerations in public–private collaborations. *Journal of Management*, 50(1):10–40.

George, G., Howard-Grenville, J., Joshi, A., and Tihanyi, L. (2016). Understanding and tackling societal grand challenges through management research. *Academy of management journal*, 59(6):1880–1895.

Ginther, D. K. and Heggeness, M. L. (2020). Administrative discretion in scientific funding: Evidence from a prestigious postdoctoral training program. *Research policy*, 49(4):103953.

Greenwood, B. N., Agarwal, R., Agarwal, R., and Gopal, A. (2019). The role of individual and organizational expertise in the adoption of new practices. *Organization science*, 30(1):191–213.

Haeussler, C. and Sauermann, H. (2020). Division of labor in collaborative knowledge production: The role of team size and interdisciplinarity. *Research Policy*, 49(6):103987.

Hoenen, S. and Kolympiris, C. (2019). Data for the value of insiders as mentors: evidence from the effects of nsf rotators on early career scientists.

Howard-Grenville, J. (2021). Caring, courage and curiosity: Reflections on our roles as scholars in organizing for a sustainable future. *Organization Theory*, 2(1):2631787721991143.

Jain, A. (2013). Learning by doing and the locus of innovative capability in biotechnology research. *Organization Science*, 24(6):1683–1700.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, 76(1):283–317.

Jones, B. F. (2021). The rise of research teams: Benefits and costs in economics. *Journal of Economic Perspectives*, 35(2):191–216.

Kaiser, U., Kongsted, H. C., Laursen, K., and Ejsing, A.-K. (2018). Experience matters: The role of academic scientist mobility for industrial innovation. *Strategic Management Journal*, 39(7):1935–1958.

Kolympiris, C., Hoenen, S., and Klein, P. G. (2019). Learning by seconding: evidence from national science foundation rotators. *Organization Science*, 30(3):528–551.

Levin, S. G. and Stephan, P. E. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American economic review*, pages 114–132.

Merton, R. K. (1968). The matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810):56–63.

Moore, T. J., Zhang, H., Anderson, G., and Alexander, G. C. (2018). Estimated costs of pivotal trials for novel therapeutic agents approved by the us food and drug administration, 2015-2016. *JAMA internal medicine*, 178(11):1451–1457.

Myers, K. (2020). The elasticity of science. *American Economic Journal: Applied Economics*, 12(4):103–134.

Nagaraj, A., Shears, E., and de Vaan, M. (2020). Improving data access democratizes and diversifies science. *Proceedings of the National Academy of Sciences*, 117(38):23490–23498.

Nagle, F. and Teodoridis, F. (2020). Jack of all trades and master of knowledge: The role of diversification in new distant knowledge integration. *Strategic Management Journal*, 41(1):55–85.

Owen, R., Pansera, M., Macnaghten, P., and Randles, S. (2021). Organisational institutionalisation of responsible innovation. *Research Policy*, 50(1):104132.

Pisano, G. P., Bohmer, R. M., and Edmondson, A. C. (2001). Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Science*, 47(6):752–768.

Polanyi, M. (1958). *Personal knowledge*. University of Chicago Press.

Polanyi, M. (1966). The logic of tacit inference. *Philosophy*, 41(155):1–18.

Qiao, Y., Alexander, G. C., and Moore, T. J. (2019). Globalization of clinical trials: Variation in estimated regional costs of pivotal trials, 2015–2016. *Clinical Trials*, 16(3):329–333.

Raelin, J. A. (1997). A model of work-based learning. *Organization science*, 8(6):563–578.

Ravetz, J. R. (1971). *Scientific knowledge and its social problems*. Oxford University Press.

Reagans, R., Argote, L., and Brooks, D. (2005). Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management science*, 51(6):869–881.

Reagans, R. and McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative science quarterly*, 48(2):240–267.

Roach, M. and Sauermann, H. (2010). A taste for science? phd scientists' academic orientation and self-selection into research careers in industry. *Research Policy*, 39(3):422–434.

Roche, M. P. (2023). Academic entrepreneurship: Entrepreneurial advisors and their advisees' outcomes. *Organization Science*, 34(2):959–986.

Sauermann, H. and Roach, M. (2012). Science phd career preferences: levels, changes, and advisor encouragement. *PloS one*, 7(5):e36307.

Senker, J. (1995). Tacit knowledge and models of innovation. *Industrial and corporate change*, 4(2):425–447.

Shibayama, S. (2019). Sustainable development of science and scientists: Academic training in life science labs. *Research Policy*, 48(3):676–692.

Simon, H. A. (1991). Bounded rationality and organizational learning. *Organization science*, 2(1):125–134.

Stephan, P. (2012). *How economics shapes science*. Harvard University Press.

Taylor-Robinson, S. D., Spearman, C. W., and Suliman, A. A. (2021). Why is there a paucity of clinical trials in africa? *QJM: An International Journal of Medicine*, 114(6):357–358.

Teodoridis, F. (2018). Understanding team knowledge production: The interrelated roles of technology and expertise. *Management Science*, 64(8):3625–3648.

Teodoridis, F., Lu, J., and Furman, J. L. (2022). Mapping the knowledge space: Exploiting unassisted machine learning tools. Technical report, National Bureau of Economic Research.

Thiers, F. A., Sinskey, A. J., Berndt, E. R., et al. (2008). Trends in the globalization of clinical trials. *Nature Reviews Drug Discovery*, 7(1):13–14.

Thompson, P. (2007). How much did the liberty shipbuilders forget? *Management science*, 53(6):908–918.

Thurstone, L. L. (1919). The learning curve equation. *Psychological Monographs*, 26(3):i.

Truffa, F. and Wong, A. (2022). Undergraduate gender diversity and direction of scientific research.

Vakili, K. and McGahan, A. M. (2016). Health care's grand challenge: Stimulating basic science on diseases that primarily afflict the poor. *Academy of Management Journal*, 59(6):1917–1939.

Vasudeva, G., Zaheer, A., and Hernandez, E. (2013). The embeddedness of networks: Institutions, structural holes, and innovativeness in the fuel cell industry. *Organization Science*, 24(3):645–663.

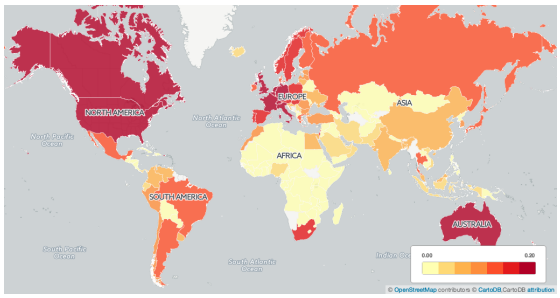
Wang, D. (2015). Activating cross-border brokerage: Interorganizational knowledge transfer through skilled return migration. *Administrative Science Quarterly*, 60(1):133–176.

Wuchty, S., Jones, B. F., and Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039.

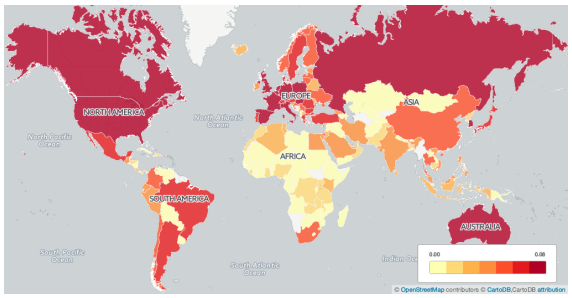
Zwick, T., Frosch, K., Hoisl, K., and Harhoff, D. (2017). The power of individual-level drivers of inventive performance. *Research policy*, 46(1):121–137.

Figure 1: Global locations of clinical trials over time

Panel A: 2000

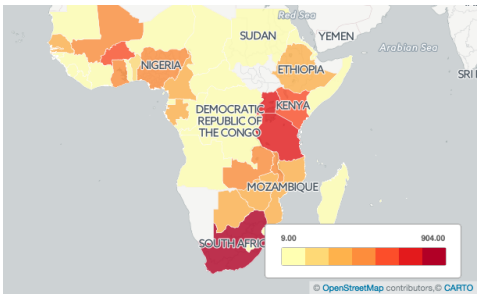


Panel B: 2015



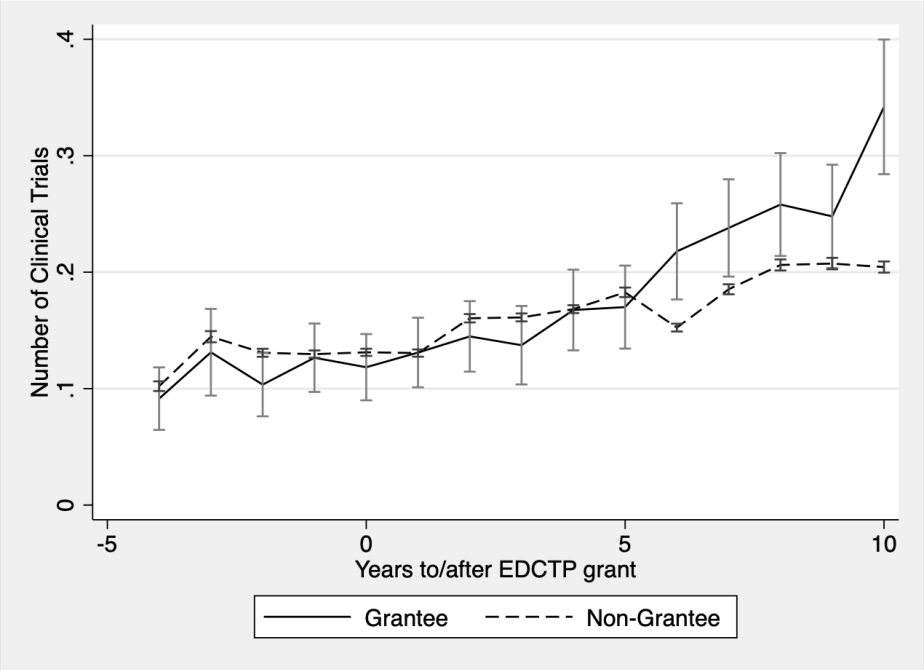
Notes: We plot the density of clinical trials indexed in ClinicalTrials.gov starting in 2000 (Panel A) and 2015 (Panel B) around the world.

Figure 2: Clinical trials in Sub-Saharan Africa



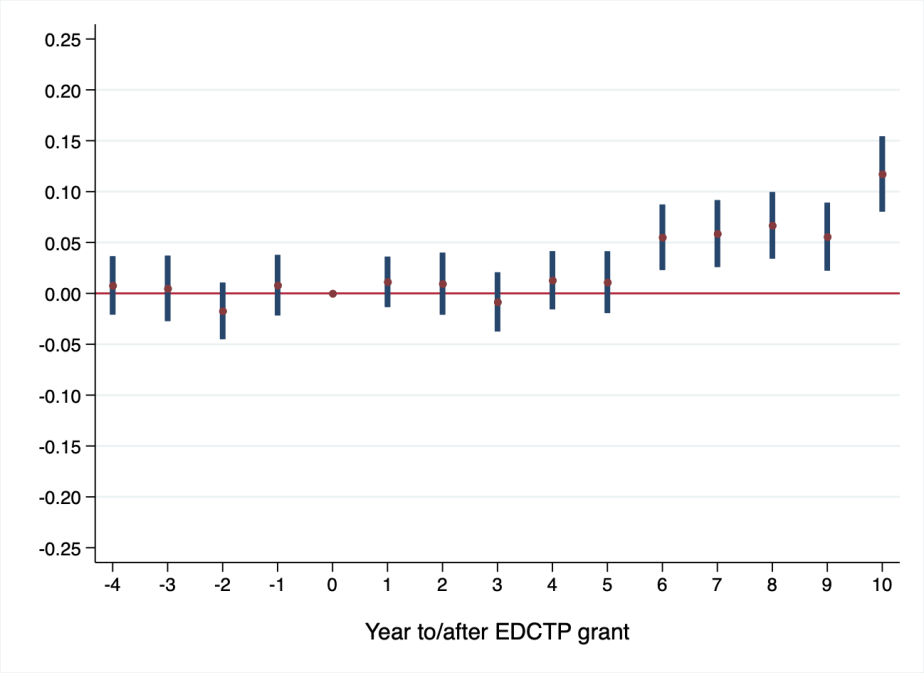
Notes: We plot the density of clinical trials indexed in ClinicalTrials.gov starting between 2000 and 2015 in countries in sub-Saharan Africa.

Figure 3: Changes in clinical trial participation after receiving an EDCTP grant



Notes: Raw trends of average clinical trials participated in per researcher in the treated and control group for the are plotted for the four years before and and ten years after the grant (or counterfactual).

Figure 4: Differences in clinical trial participation between treated and untreated scientists



Notes: Coefficient estimates stemming from conditional (scientist) fixed effects ordinary least squares specifications in which inverse hyperbolic sine clinical trials are regressed onto year effects, scientist age effects, as well as interaction terms between treatment status and the number of years before/after the EDCTP grant (or counterfactual). The 90% confidence interval robust standard errors clustered around the institution is plotted with solid bars.

Table 1: Differences in pre-treatment variables between EDCTP and matched control scientists

	Treated Scientists	Control Scientists	Difference in Means	P-Value of Differences
Number of Clinical Trials	0.329	0.358	-0.030	0.526
Number of Trials at Institution	43.288	43.652	-0.39	0.935
Size of Institution	227.78	248.95	-21.16	0.305
Number of OECD Co-authors	6.600	6.365	0.235	0.668
Number of Publications	6.246	5.596	0.650	0.012
Number of HIV/TB/Malaria Publications	2.257	1.423	0.840	0.000
Career Age in Treatment Year	9.208	9.009	0.198	0.544
Number of Scientists	794	144,475		

Notes: Differences between treated and matched control scientists during the three-year period before the treated scientists first receive the EDCTP grant (and counterfactual control grant). Career age is measured in the final year before a scientist's treatment event.

Table 2: Changes in clinical trial participation after receiving an EDCTP grant

	Number of Clinical Trials		EDCTP Disease Trials	Non-EDCTP Disease Trials	SNIP Weighted Trial Publications	First or Last Authored Trial Publications
	(1)	(2)	(3)	(4)	(5)	(6)
EDCTP Grantee × Post-Grant	0.0365***	0.0339***	0.0234***	0.0169**	0.0272***	0.0060*
	(0.010)	(0.010)	(0.005)	(0.008)	(0.008)	(0.004)
Total Observations	1832217	1832217	1832217	1832217	1832217	1832217
Mean of Dep. Variable	0.0179	0.0179	0.0018	0.0161	0.0067	0.0031
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs		X				

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts of non-EDCTP clinical trials.
[b] Heteroskedastic robust standard errors are given in parentheses.

Table 3: Clinical trial sponsorship

	EU Sponsored Clinical Trials	US Sponsored Clinical Trials	Africa Sponsored Clinical Trials
	(1)	(2)	(3)
EDCTP Grantee × Post-Grant	0.0056*	0.0228***	0.0043
	(0.003)	(0.007)	(0.003)
Total Observations	1832217	1832217	1832217
Mean of Dep. Variable	0.0022	0.0075	0.0014
Author FE	X	X	X
Year FE	X	X	X
Career Age FE	X	X	X
Time Since Grant FE	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts.
[b] Heteroskedastic robust standard errors are given in parentheses.

Table 4: Changes in scientists' research direction after receiving an EDCTP grant

	Number of Publications on Diseases	Number of Non-Disease Publications	Number of Publications Trial Disease	Number of Publications Non-Trial Disease	Number Basic Publications	Number Applied Publications
	(1)	(2)	(3)	(4)	(5)	(6)
EDCTP Grantee × Post-Grant	0.1277***	-0.0555***	0.1768***	-0.0061	-0.0285*	0.0722***
	(0.014)	(0.014)	(0.019)	(0.017)	(0.015)	(0.014)
Total Observations	1832217	1832217	1832217	1832217	1832217	1832217
Mean of Dep. Variable	0.3266	0.5489	0.0836	0.2601	0.3713	0.2892
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs	X	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts.
[b] Heteroskedastic robust standard errors are given in parentheses.

Table 5: Heterogeneity in changes in scientists' clinical trial participation by prior trial experience

	Number of Clinical Trials		EDCTP Disease Trials	Non-EDCTP Disease Trials	SNIP Weighted Trial Publications	First or Last Authored Trial Publications
	(1)	(2)	(3)	(4)	(5)	(6)
EDCTP Grantee × Post-Grant						
× Prior Trial	-0.0687**	-0.0688**	-0.0580***	-0.0018	0.0010	-0.0022
	(0.034)	(0.034)	(0.016)	(0.031)	(0.030)	(0.013)
EDCTP Grantee × Post-Grant	0.0524***	0.0496***	0.0365***	0.0177***	0.0267***	0.0066**
	(0.008)	(0.008)	(0.005)	(0.006)	(0.006)	(0.003)
Total Observations	1832217	1832217	1832217	1832217	1832217	1832217
Mean of Dep. Variable	0.0179	0.0179	0.0018	0.0161	0.0067	0.0031
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs		X				

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. The variable Prior Trial takes the value of 1 if a focal scientist has participated in a clinical trial in the 3 years prior to the EDCTP trial, 0 otherwise. [b] Heteroskedastic robust standard errors are given in parentheses.

Table 6: Heterogeneity in changes in scientists' clinical trial participation by career age

	Number of Clinical Trials		EDCTP Disease Trials	Non-EDCTP Disease Trials	SNIP Weighted Trial Publications	First or Last Authored Trial Publications
	(1)	(2)	(3)	(4)	(5)	(6)
EDCTP Grantee × Post-Grant × Senior Scientist	-0.0301*	-0.0291*	0.0056	-0.0344**	-0.0087	-0.0101
	(0.017)	(0.017)	(0.010)	(0.014)	(0.013)	(0.006)
EDCTP Grantee × Post-Grant	0.0566***	0.0534***	0.0196***	0.0399***	0.0330***	0.0127***
	(0.011)	(0.011)	(0.007)	(0.009)	(0.008)	(0.004)
Total Observations	1832217	1832217	1832217	1832217	1832217	1832217
Mean of Dep. Variable	0.0179	0.0179	0.0018	0.0161	0.0067	0.0031
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs		X				

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. The variable Senior takes the value of 1 if the focal scientist has more than 5 years since their first publication at the time of the EDCTP trial. [b] Heteroskedastic robust standard errors are given in parentheses.

Table 7: Alternative explanations for heterogeneous changes by prior trials

	Number of Clinical Trials					
	(1)	(2)	(3)	(4)	(5)	(6)
EDCTP Grantee × Post-Grant × Prior Country Trials	-0.0687**	-0.0625*	-0.0677*	-0.0656*	-0.0594*	-0.0653*
	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)	(0.035)
EDCTP Grantee × Post-Grant	0.0524***	0.0769***	0.0484***	0.0502***	0.0503***	0.0558***
	(0.008)	(0.013)	(0.009)	(0.009)	(0.010)	(0.010)
EDCTP Grantee × Post-Grant × Career Age		-0.0025**				
		(0.001)				
EDCTP Grantee × Post-Grant × JIF Pubs			-0.0001			
			(0.001)			
EDCTP Grantee × Post-Grant × OECD Pubs				0.0008		
				(0.003)		
EDCTP Grantee × Post-Grant × Applied Pubs					-0.0016	
					(0.004)	
EDCTP Grantee × Post-Grant × HIV, Malaria, TB Pubs						-0.0028
						(0.004)
Total Observations	1832217	1832217	1832217	1832217	1832217	1832217
Mean of Dep. Variable	0.0179	0.0179	0.0179	0.0179	0.0179	0.0179
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs						

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts.
[b] Heteroskedastic robust standard errors are given in parentheses.

Table 8: Changes in research output and collaborations by prior trial experience

	Number of Publications	Impact Factor Weighted Publications	First or Last Authored Publications	Number of OECD Coauthored Publications	Number of Non-OECD Coauthored Publications
	(1)	(2)	(3)	(4)	(5)
EDCTP Grantee × Post-Grant × Prior Trial	0.0148	0.0814	0.1072**	0.1698**	-0.0163
	(0.082)	(0.098)	(0.047)	(0.073)	(0.068)
EDCTP Grantee × Post-Grant	0.5497***	0.6005***	0.0783***	0.3461***	0.3932***
	(0.032)	(0.036)	(0.014)	(0.027)	(0.027)
Total Observations	1832217	1832217	1832217	1832217	1832217
Mean of Dep. Variable	0.8755	0.8663	0.1882	0.3350	0.5445
Author FE	X	X	X	X	X
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. The variable Prior Trial takes the value of 1 if a focal scientist has participated in a clinical trial in the 3 years prior to the EDCTP trial, 0 otherwise.
[b] Heteroskedastic robust standard errors are given in parentheses.

Table 9: Use and reuse of new knowledge and skills developed in EDCTP trial, by prior trial experience

	Number of New-to-the-Scientist Terms in Trial Pubs	Number of Repeated New-to-the-Scientist Trial Terms	Number of Repeated New-to-the-Scientist Trial Terms	Number of Repeated Previously Indexed Trial Terms	Number of Repeated Previously Indexed Trial Terms
	(1)	(2)	(3)	(4)	(5)
EDCTP Grantee × Prior Trial	-0.3695***	-0.0794	-0.0873*	0.3339***	0.3376***
	(0.057)	(0.048)	(0.047)	(0.075)	(0.073)
Total Observations	1218	26653	26653	26653	26653
Mean of Dep. Variable	0.3478	0.9003	0.9003	2.7748	2.7748
Year FE	X	X	X	X	X
Career Age FE	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X
Terms in Trial Pub	X				
Total Trial Terms		X	X	X	X
Total New-to-Scientist Trial Terms			X		
Total Previously Indexed Trial Terms					X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from ordinary least squares regression specifications at the publication level. In Model 1, the dependent variable is the inverse hyperbolic sine transformed count of MeSH terms indexed to a publication linked to an EDCTP trial that have not previously been indexed to a scientist's publications prior to the trial. We control for the total number of terms indexed to that trial-linked publication. In Models 2 to 5, the dependent variables are inverse hyperbolic sine transformed counts of terms indexed to a scientist's publications measured at the publication level. These are defined according to whether a term from a publication linked to an EDCTP trial was new-to-the-scientist (columns 2 and 3) or had previously been indexed to a scientist's publications prior to their participation in an EDCTP trial (columns 4 and 5). We control for the number of terms indexed to the scientist's papers across all their trial-linked publications.
[b] Sample in Model 1 only includes publications from trials linked to EDCTP-sponsored trials and includes only treated scientists. The sample in other Models includes all publications among treated scientists from the first year of participation in the EDCTP trial onwards. In Model 3, we control for the number of new-to-the-scientist terms that were indexed to their EDCTP trial publications. In Model 5, we control for the number of previously-indexed terms that were indexed to their EDCTP trial publications. Only publications with indexed MeSH terms in PubMed are included in the samples.
[c] Heteroskedastic robust standard errors are given in parentheses.

Table 10: Heterogeneity in changes in scientists' research trajectory by country level prior trials

	Number of Clinical Trials Low Capacity Country		Number of Clinical Trials High Capacity Country			
	(1)	(2)	(3)	(4)	(5)	(6)
EDCTP Grantee × Post-Grant × Prior Trial		-0.1473**	-0.1473**		-0.0402	-0.0404
EDCTP Grantee × Post-Grant	0.0168	0.0528***	0.0519***	0.0418***	0.0497***	0.0459***
	(0.018)	(0.014)	(0.014)	(0.012)	(0.009)	(0.010)
		(0.060)	(0.060)		(0.041)	(0.041)
Total Observations	816371	816371	816371	1015844	1015844	1015844
Mean of Dep. Variable	0.0152	0.0152	0.0152	0.0201	0.0201	0.0201
Author FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X
Annual Pubs			X			X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares regression specifications in which dependent variables are inverse hyperbolic sine transformed annual counts. The variable Prior Country Trials takes the value of 1 if a focal scientist is affiliated with a country which is above the median sample value in terms of the number of clinical trials taking place in that country in the 3 years prior to the EDCTP trial, 0 otherwise. [b] Heteroskedastic robust standard errors are given in parentheses.