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

November 10, 2023

Abstract

In this study we ask whether a scientist's experience gained through participating in a publicly funded project spills over to follow on innovative projects, particularly in resource poor environments. We investigate this question in the context of participation in clinical trial research in sub-Saharan Africa. We document that participation in a publicly funded clinical trial increases scientists' future involvement in trials and in more applied-disease focused research, with the most significant impact seen in scientists without prior trial experience and those with fewer years of scientific experience. This suggests that knowledge and skills acquired in the initial project are transferable to subsequent endeavors, and that exposure to specific types of projects are most impactful for early career scientists with less experience initially. In other words, early experience plays a significant role in the research path that scientists' pursue longer term. The study also highlights the importance of international cooperation among funding organizations to enhance scientific capacity in resource-poor environments, particularly in the context of clinical trials in sub-Saharan Africa. It underscores the potential for spillovers between funders' projects benefiting global scientific development and health outcomes in regions with limited resources.

1 Introduction

In 2021, the world's first malaria vaccine, Mosquirix, was approved for global use by the WHO. The development for this vaccine started in 1987, with phase 3 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 3 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." This anecdote raises the question: How does a scientist's experience in a particular innovative process shape their involvement in subsequent innovative activities, particularly early in their career?

A large literature has explored the magnitude and impact of knowledge spillovers from public investments in science and innovation (Nelson, 1959; Griliches, 1992). Knowledge spillovers from public investment in science has been shown to increase the productivity of local firms (Jaffe, 1989; Zucker and Darby, 1997; Furman et al., 2006), to increase private sector research productivity within a scientific field (Li et al., 2017; Azoulay et al., 2019; Moretti et al., 2021; Howell, 2017; Myers and Lanahan, 2021), and to induce other actors to shift their investments in science and innovation (Kyle et al., 2017). Researchers have also documented inter-organizational spillovers from academic science to other organizations through mechanisms such as collaborative research and labor mobility (Zucker et al., 2002; Furman and MacGarvie, 2007) or more direct knowledge transfer through scientific publications (Bryan and Ozcan, 2021). However, we know relatively less about how a scientist's experience gained through participating in a publicly funded project spills over to their own follow on innovative projects, particularly in resource poor environments. As resources are provided to scientists for distinct projects, there may be spillovers between projects as scientists develop knowledge and skills in one project that that can be exploited in their subsequent research.

Participation in an innovative process could increase the rate of subsequent involvement in similar activities because of improved proficiency in necessary tasks as scientists gain more experience and enhance their knowledge and skills. In resource poor environments where there are fewer opportunities to participate in these processes, the marginal gains in proficiency from project experiences may be particularly great as scientists have fewer other opportunities to develop relevant research skills. But, there may also be limits to the returns to experience in resource poor environments because subsequent innovation often relies on attracting external resources. This is a task which itself may be difficult, especially for those scientists who have fewer existing resources to leverage and who are more distant from hubs of global science. Therefore, a key question is: Does experience in a publicly funded project increase the rate at which a scientist in a resource poor environment participates in subsequent, similar projects?

Experience may not only affect the average rate at which scientists participate in subsequent, similar

projects. It is also plausible that there are heterogeneous effects across scientists, which could affect the longer-term distribution of resources and performance across scientists. Therefore, a second important question is: Which scientists are most impacted by experience in publicly funded projects? On the one hand, according to the principles of cumulative advantage, we might expect that scientists with prior experience in similar projects, or at a later stage in their career, are more likely to be positively affected by project experience. They may have complementary knowledge, networks, or resources to draw upon that can be used to leverage the additional experience to participate in future projects. On the other hand, however, there may be diminishing returns to learning in resource poor environments. Such diminishing returns could result from limits to the opportunities available, saturation of knowledge related to tasks in high demand in resource poor environments, or later career scientists having capacity constraints or preferences for diversifying their research. In this case, we might expect that scientists with less prior experience in similar projects, or earlier career scientists, are more likely to benefit.

We explore these questions in the context of publicly funded clinical trials in sub-Saharan Africa. This is an important context to study these questions. More efficient testing and development of potential drugs and vaccines that target diseases with a large burden in the Global South, such as malaria, TB and HIV, and Sub-Saharan Africa in particular, would have enormous impact on the lives of those affected by these diseases. Yet, despite its disproportionate burden of global disease, relatively few clinical trials take place in Sub-Saharan Africa. Clinical trials are an essential, but enormously expensive and lengthy, part of the drug development process. By their very nature, the costs and feasibility of trials are dependent on pre-existing institutions and investigator competency. Because of the varying disease profiles of diseases around the world, the location of clinical trials for drugs and vaccines for these diseases is somewhat fixed (particularly for larger phase 2 and 3 trials examining clinical effectiveness).

Currently, the volume of clinical trials taking place in sub-Saharan Africa, in particular, is extremely low. Trials are very costly, and developing the institutional capabilities and human capital to run clinical trials involves significant fixed investments and training costs for investigators. Additionally, a wide range of public and private organizations fund different trials. If knowledge and skills developed by scientists in one funder's project transfer to other projects, such spillovers may not be fully captured by a single funder. If the social returns to developing trial capabilities thus exceed the returns to any single funder, subsidies for developing these capabilities may help increase the number clinical trials taking place in Sub-Saharan Africa. In resource poor environments with relatively fewer opportunities for scientists to develop these knowledge and skills, such policies may have a particularly large impact on scientists' ability to participate in future clinical trials.

We examine one such program in this paper. The European and Developing Countries Clinical Trial Partnership (EDCTP) is an EU funded partnership between 15 European countries that has funded 140 late-stage clinical trials in HIV, malaria, and tuberculosis in Africa to date. They offer funding at the trial level, but also tag funding for individual scientists' for career development and masters and PhD trial par-

ticipation. We evaluate the effect of participating in an EDCTP funded trial between 2003 and 2015. To do so, we construct a panel dataset of 908 African scientists who participated in an EDCTP trial. Matching with scientists who are not participating in an EDCTP trial, we are able to control for career, field, and temporal changes in research output and clinical research participation. Difference-in-differences regressions compare within scientist changes in clinical trial involvement and research output of scientists following the participation in an EDCTP funded trial with changes of observably similar scientists who were not involved in an EDCTP funded trial. Variation according to career stage and clinical trial experience before an EDCTP funded trial is explored.

The results reveal that participation in an EDCTP funded trial increases scientists' future participation in clinical trials and increases the proportion of publications devoted to more applied, disease focused topics. In terms of magnitude, treated scientists participate in around 3.5% more clinical trials following the EDCTP trial, which is approximately a 300% increase on the mean. Interestingly, they do not increase the rate at which they participate in industry sponsored trials or African sponsored trials, but they do tend to be involved in both more EU and US sponsored trials, implying that there are some spillovers across government funders. This increase is most pronounced for scientists without prior experience in clinical trials, and for earlier career scientists, suggesting that first participation in clinical trials matters the most. In fact, scientists who have previously worked on clinical trials show little evidence of increased participation in future trials relative to untreated control scientists with similar prior clinical 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.

We use bibliographic data to provide indicative evidence of this learning mechanism. We identify the MeSH terms that are independently indexed to scientists' research by the US National Library of Medicine. We find that the trial-linked publications of scientists without prior trial experience includes a greater share of indexed keywords that have not previously been indexed to the scientist's research, compared to those with prior trial experience. We also show that the subsequent research of scientists without prior trial experience is less likely to be associated with MeSH terms that were indexed to their research before the focal trial compared to scientists with prior trial experience. This suggests that there is a shift in the direction of scientists' research towards projects that utilize the knowledge and skills developed through participation in a first clinical trial.

Taken together, our results imply that the knowledge and skills scientists develop by participating in an initial clinical trial transfer to subsequent trials. This indicates that there are knowledge spillovers from a first funder's projects to subsequent projects carried out by a scientist. Individual learning takes place in the first project, which is applied in future projects. Since these returns are not entirely captured by the initial funder, the social return to funding scientists' participation in a first clinical trial may exceed the private returns to that funding organization. This highlights a public goods problem in the development of global

scientific capacity in resource poor environments. There are many different funders, from many different countries, none of whom capture the full returns to developing this scientific capacity. Our results point to an important role for international co-operation among funding organizations to invest in scientific capacity in resource poor countries where this is necessary to advance the developments of drugs and vaccines to improve health outcomes in those regions.

We attempt to make several contributions in this study. First, our results highlight a channel for spillovers from publicly funded innovation and science: individual learning. This raises the question of which funder should be responsible for bearing the first (and largest) costs, particularly with distributed global innovative activities in resource poor environments. Second, we show that experience can have a distributional effect on innovation. According to the principles of accumulative advantage, we might expect that those with more prior experience leverage the trial to participate in future trials. However, we find the opposite. This implies a focus on earlier career scientists and those with less experience could have outsized gains. Third, this study contributes to the work on the drivers of innovation. Extant literature has focused predominantly on demand side drivers (Kremer and Glennerster, 2004; Kyle and McGahan, 2012), particularly when it comes to innovation for diseases that affect the world's poorest. Our results show that the provision of an opportunity to gain experience and learning of the clinical trial process to scientists in resource poor environments can also act as a driver to propel innovation for some of the world's grandest challenges.

2 Setting

2.1 The Global Distribution of Clinical Trials

2.1.1 Drug and vaccine development and clinical trials

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 1 trials test for safety and tolerability of the drug in healthy volunteers, Phase 2 trials assess preliminary efficacy of the drug, while Phase 3 clinical trials assess the safety and efficacy in a larger population, and are often carried out at several sites, spanning the globe. Lastly, Phase 4 trials typically take place after regulatory approval and aim to provide evidence on longer-term side effects.

Clinical trials from phase two 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. In addition to the costs, with fixed term patent

lifetimes starting prior to initiation of the clinical trials, the length of the trial also impacts the ultimate return on investment from the product (Budish et al., 2015).

2.1.2 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, moving many of their trials recently to Eastern Europe and Asia. 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 diseases, 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).

[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.

2.2 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 policy related infectious diseases in sub-Saharan Africa, through all phases of clinical trials, with an emphasis on Phase 2 and 3 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. The aim was to pool European resources to make progress to combat these diseases. With a budget of over 1 billion euros, EDCTP-1 was the largest program to support clinical trials in sub-Saharan Africa to date.

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 a trial.

One such sponsored trial was a trial known as TaMoVac-01, which was a Phase 1/2 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.

An example fellowship project is a senior fellowship of Abdoulaye Djimde who is based in the Malaria Research and Training Center in Mali. This project started in 2005, finished in 2009 and the project itself centered around a Phase 4 randomized clinical trial that tested the effect of repeated administration of three different malaria drugs (AS+AQ, AS/SP and AR-L) on malaria immunity and transmission in 780 individuals.

3 Empirical Strategy and Data

3.1 Empirical Strategy

In this paper we estimate the effect of being involved in an EDCTP sponsored clinical trial (via either project based or fellowship funding mechanisms). 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. Our strategy allows us to account for individual level heterogeneity, career age trends and year trends through the use of fixed effects. 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.

3.2 Sample construction

3.2.1 EDCTP trial participants - treated scientists

We identified 1,215 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 1,215 scientists, 444 participated in HIV trials, 368 in malaria trials, and 433 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, 21 percent were from South Africa, 10 percent from Tanzania, 10 percent from Uganda and 9 percent from Kenya. Overall, 908 of the 1,215 scientists had a publication record prior to the trial.

3.2.2 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 823 treated scientists, and 144,475 control scientists.

3.3 Variables and Measures

The data in our sample is constructed from three main sources. We draw on clinical trials data from ClincialTrials.gov, publications data from Scopus, and link publications in Scopus to the keywords indexed to papers in the PubMed database to add additional richness to our data. The ClincialTrials.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.[1] 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 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 augment this data by building a rich bibliographic dataset based on scientists' publications. We use the Scopus author identifier to extract each sample scientist's publications recorded in this database. We use the bibliographic data in Scopus to identify the diseases of focus in scientists' papers, based on keywords in their abstracts, scientists' institutional affiliations over their careers, numbers of citations to their papers, and the affiliations of their co-authors. We also use the PubMed identifiers of papers recorded in Scopus to link these publications to MeSH terms, or descriptor terms used to organize concepts in research.[2] Each paper in our sample is placed at a particular point in the space of scientific concepts based on its content on the dimensions measured by each MeSH term. 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.

[1] 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.

[2] 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 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.

3.3.1 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). As a robustness check, we first exclude all trials linked to the EDCTP program. Second, to examine whether our results are specific to particular types of clinical trials, we separately construct the dependent variable based on the whether a clinical trial is interventional or observational. We then repeat the analysis with dependent variables that are defined according to the phase of a focal trial (i.e., Phase 1 to Phase 4). 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. Finally, 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 EDCTP program—namely, HIV, malaria, or tuberculosis—or a different disease. This enables us to examine how far experience in an EDCTP-supported trial is linked to scientists participating in clinical trials for other types of disease.

Second, we create a set of dependent variables to track both quantitative and qualitative 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, 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. The specific definitions are presented alongside the models in Section 4 below. 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). A publication is defined as basic if it appears in a journal with a zero JCIF. We then create separate variables based on the counts of scientist i's applied and basic publications in year t. 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. For robustness, we then create

two similar variables based on the number of scientist i's co-authors in year t that are 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 trial by the EDCTP. We then create a variable that measures the share of MeSH terms indexed to a trial-linked publication by scientist i, which had not been previously 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 a clinical trial. Second, we create a dependent variable that measures the number of MeSH terms on a post-treatment publication by scientist i that had been indexed to scientist i's publications in years prior to treatment.

3.3.2 Descriptive statistics

Table 1 presents any differences in clinical trial participation, publications, institutions, co-authorship relationships, and experience between treated and control scientists prior to the EDCTP trial. We show the difference in means of variables that are constructed for a three-year period prior to the year in which the treated scientists in a given stratum first participated in the EDCPT program.

We can see that 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. 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.

[Table 1 about here.]

4 Results

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

Our core regression model 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 Scopus 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. This serves to 'line up' each scientist's pre- and post-treatment periods since the treatment years can vary between the different strata of treated and control scientists depending on when the treated scientist received an EDCTP grant.

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

$$y_{it} = Post_{kt} * Grantee_i + s_i + t_t + c_{it} + d_{kt} + \varepsilon_{ikt} \tag{1}$$

The variable *Post* is denotes whether the scientists in the coarsened exact matching stratum k are in the pre- or post-treatment period for that stratum. The variable *Grantee* is our treatment variable. 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_{kt} represents the time to/from treatment fixed effect for the scientists in stratum k (which also absorbs the main effect of $Post_{kt}$ in the model). Standard errors are clustered at the scientist level.

4.1 Participation in Clinical Trials

Our first set of results concerns the impact of participating in the EDCTP program on 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 African scientists participate after EDCTP treatment. If we interpret the point estimate of the coefficient, 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 Column 3, we exclude trials linked to an EDCTP grant from the construction of our dependent variable. The results are almost identical to those in Column 1, implying that our results our not driven by an increased participation in trials directly linked to the EDCTP treatment grant.

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. The trends shown in Figures 3 and 4 are, for example, consistent with EDCTP participants gaining experience during a clinical trial that supports future trial participation or completing a trial providing a signal of competence that facilitates further participation in trials.

[Figure 3 about here.]

[Figure 4 about here.]

Next, we break down the results by type of clinical trial. In Columns 4 to 9 of Table 2, we define the dependent variable according to whether a trial is interventional or observational and according to the phase of the drug development process with which the trial is associated. We find that there is an increase in treated scientists' participation in both interventional and observational trials. However, although positive, the effects of receiving an EDCTP grant on subsequent participation in Phase 2 or 3 clinical trials are not significant.

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 industry or non-industry funders, with sponsors from the EU, US, or Africa, and on trials for diseases in the EDCTP grant program or for other diseases. The results in Columns 1 and 2 indicate that the increase in post-treatment participation in clinical trials is driven exclusively by trials that are not sponsored by industry. Treated scientists are no more likely than their matched controls to participate in industry-sponsored trials after participating in the EDCTP program. However, industry-sponsored trials are relatively rare in our data, amounting to fewer than 10% of all clinical trials.

[Table 3 about here.]

In Columns 3 to 5 of Table 3, we examine changes in scientists' participation in EU-, US-, or Africa-sponsored clinical trials after participating in the EDCTP program. There are significant increases in treated scientists' participation in clinical trials sponsored by EU and US organizations. The EDCTP program provided grants for working on clinical trials with European partners. Thus, these results provide evidence that treated scientists increase the rate at which they participate in clinical trials sponsored by organizations with which they had not previously worked in the EDCTP program. This indicates that valuable knowledge and skills scientists developed in the EDCTP program may be spilling over to other funders' projects, rather than being specific to a relationship with a given funder. However, we find little conclusive evidence of changes in scientists' rate of participation in Africa-sponsored clinical trials after the EDCTP program.

While the point estimate of the coefficient is positive and large relative to the mean of the dependent variable, African-sponsored trials are rare and the coefficient is imprecisely estimated with a large standard error.

We next examine how participation in the EDCTP program affects participation in subsequent clinical trials across different diseases. EDCTP grants were focused on trials for drugs and vaccines for HIV, malaria, and tuberculosis. If treated scientists are subsequently more likely to work on clinical trials in other disease areas, this would suggest that the knowledge and skills developed by participating in an EDCTP trial can be transferred to projects in other disease areas. This would indicate that spillovers may take place across projects of different diseases, as well as across funding organizations. The results in Columns 6 and 7 of Table 3 show that this is the case. Treated scientists are relatively more likely to participate in subsequent clinical trials for both EDCTP-funded diseases and other diseases. While the proportional increase is greater for EDCTP-funded diseases, this indicates that relevant learning is not all disease-specific and may be applicable to clinical trials for other diseases.

Finally, in Columns 8 and 9, we present two additional results. First, 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 citations to trial publications (weighted by source normalized impact per paper). The relative increase in the rate of citations to trial publications following EDCTP participation is greater than the increase in the rate of trial participation. This suggests that treated scientists' increase in trial participation is not driven by low quality trials that have less impact on the scientific community. Second, we show that there is not a significant change in the rate at which treated scientists are the first or last author on trial publications. The relevant coefficient is positive, but not statistically significant. A null result would be consistent with treated scientists being no more likely to take on leadership roles in trials after EDCTP participation, although we are hesitant to make inferences based on this result.

4.2 Changes in Research Direction

We now examine other changes in scientists' research after participating in the EDCTP program. In particular, we are interested to see whether there are changes in scientists' research that suggests that they are changing their research direction after EDCTP participation. 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. In Table 4, we use publications data to show that treated scientists have significant changes in their research direction after EDCTP participation. The dependent variables count the number of scientists' publications in a given year that have specific topics based on diseases and type of research project. We control for the number of publications a scientist has in a given year in these models. This is so we can isolate changes in the direction of scientific activity distinct from changes in its rate.

[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. 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 the EDCTP trial diseases—HIV, malaria, and tuberculosis—with no change on publications focused on other diseases. Next, in Columns 5 and 6, we provide further evidence that this is linked to participation in the EDCTP program. First, we showing that there is a larger increase for treated scientists in publications focused on the disease linked to a grantee's EDCTP trial. Then we show that treated scientists in fact shift their research away from the other EDCTP diseases that were not the subject of their grant. Finally, the results in Columns 7 and 8 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.

Taken together, the results in Table 4 imply that the EDCTP program led African scientists re-direct their relative allocation of research effort towards the diseases for which they received the grant to participate in a clinical trial. Treated scientists also increase their engagement in applied research relative to basic research. This is consistent with scientists either developing knowledge and skills in these areas or becoming better connected to future research opportunities through participation in the clinical trial, leading them to continue to focus on other research projects in these areas.

4.3 Heterogeneity by Prior Trial Participation and Career Age

Participation in a first clinical trial might be especially important to scientists if there are strong returns to learning-by-doing in clinical trials. In this case, participating in a trial would help scientists develop knowledge and skills that they can use in future clinical trials. 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 treatment effect 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 the date at which treated scientists in their stratum of matches received an EDCTP grant. 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:

$$y_{it} = Post_{kt} * Grantee_i * PriorTrial_i + Post_{kt} * Grantee_i + Post_{kt} * PriorTrial_i + s_i + t_t + c_{it} + d_{kt} + \varepsilon_{ikt} \quad (2)$$

The results in Tables 5 and 6 replicate those from Table 2 using our triple interaction model. There

is a clear pattern in which the treatment effects 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. Similarly, those with more experience more generally have a smaller effect from participating in the EDCTP program. Notably, the estimates in Column 1 of Table 5 suggest that these scientists had no observable increase in subsequent trial participation after the EDCTP grant. However, for scientists without prior trial experience, participation in the EDCTP program was 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.]

4.4 Individual vs. Institutional Effects

The results presented so far in this paper have shown that African scientists who participated in a clinical trial through the EDCTP program were more likely to be involved in subsequent clinical trials—driven by those without prior trial experience. However, trial experience might build capabilities to participate in future clinical trials at either the individual or institutional level. On the one hand, having a prior trial experience may matter because it provides scientists with the opportunity to learn-by-doing and develop knowledge and skills that facilitate their future trial participation with other partners. On the other hand, scientists may have participated in a prior trial because their institution has developed capabilities at the organization-level that facilitates its scientists working on clinical trials.

To separate the potential individual and institutional effects, we replicate the analysis from Table 5 but define prior trial at the organization level. Specifically, we replace the individual-level *PriorTrial* variable with one that denotes whether the organization for which scientist i works has had other scientists participate in a trial in the ClincialTrials.gov database in a year prior to i's first participation in the EDCTP program. By contrasting the results of this analysis, presented in Table 7, with those in Table 5, we can see whether individual or institutional clinical trial experience is associated with the heterogeneous response to participation in the EDCTP program. If we were to find a smaller treatment effect for EDCTP grantees in organizations with previous clinical trial experience, this would be consistent with there being institution-level trial capabilities that can be leveraged by scientists without individual-level trial experience. Consequently, scientists in these organizations should experience a lower treatment effect than scientists in other organizations that lack this institution-level experience.

[Table 7 about here.]

The results in Table 7 provide little evidence of an institution-level effect of an organization's prior trial experience on scientists' future participation in clinical trials. The results contrast with those of Table 5 in

which we show that prior trial experience offsets the main effect of EDCTP participation on future trial participation. Further evidence for the lack of an institution-level effect is provided in Figure 5. Here, we contrast the trends in clinical trial participation among untreated scientists who were either in an institution in which a colleague received an EDCTP grant or an institution in which no scientists received an EDCTP grant. Figure 5 shows that there is no evidence of a spillover effect to colleagues from an institution being home to an EDCTP grantee.

[Figure 5 about here.]

Taken together, these findings are consistent with a mechanism in which individual scientists develop knowledge and skills through trial participation that affects their future participation in clinical trials. They contrast with an institution-level mechanism in which clinical trials within an organization build capacity that other scientists—without individual trial experience—can leverage to participate in clinical trials. The results suggest that to build clinical trial capabilities in Sub-Saharan Africa, it is crucial to provide opportunities to participate in clinical trials at the individual, not just institutional, level. Thus, any spillovers between different funders' projects would thus appear to be taking place within individual scientists, rather than spilling over across projects carried out by different scientists in the same organization.

4.5 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.]

The results in Table 8 are based on the same triple interaction models as those in Section 4.3. 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 the visibility of their research 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 and citations received. This suggests that receiving and EDCTP grant is not associated with a differential rate of increase in the productivity or recognition received by first-time trial participants compared to repeat trial participants. 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 to 8 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. Both groups increase these collaborations to a similar extent. This suggests that first time trial scientists are not expanding their number of collaborators relatively more than repeat trialists, after participating in the EDCTP program, such that an increased network provides greater opportunities for future trial participation. Indeed, the repeat trialists in the EDCTP program have an (insignificantly) larger point estimate for the rate of collaboration with scientists in OECD countries, where most trial sponsoring organizations are located.

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 papers by the National Library of Medicine 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. To examine this, we carry out pooled regressions and restrict our analysis to publications directly linked to EDCTP trials.

[Table 9 about here.]

The results in Column 1 of Table 9 show that a lower share of MeSH terms indexed to repeat trialists' EDCTP trial-linked publications had not been previously linked to the scientist's research in prior years. On average, approximately 35% of MeSH terms in EDCTP trial-linked publications are new-to-the-scientist. However, the share of new-to-the-scientist terms among repeat trialists' publications is 11 percentage points lower than for first time trialists. This provides evidence that a first trial is associated with a greater proportional extension of scientists' knowledge and skills. Conversely, those with prior trial experience appear to

re-use existing knowledge and skills at a higher rate. In Columns 2 to 4, we extend this analysis to scientists' post-EDCTP publications. We show that the rate at which repeat trialists' subsequent research re-uses their pre-EDCTP knowledge and skills—as measured by the MeSH terms indexed to their publications—is higher than for first time trialists. This is consistent with repeat trialists having a greater range of knowledge and skills before participation in the EDCTP program that is relevant to their ongoing research, whereas first time trialists developed new knowledge and skills that is re-used in their future projects. The results are similar when we control for the number of terms that had been indexed to a scientist's papers before participating in the EDCTP trial. This implies that the result is not driven by repeat trialists having a larger vocabulary of indexed MeSH terms ex ante, which can then be repeated.

Overall, the results in Table 9 are consistent with a learning-by-doing mechanism. We find evidence that first participation in a clinical trial is associated with scientists building new knowledge and skills at a higher rate than repeat trialists. This suggests that knowledge may spillover between different projects. Scientists appear to learn through participating in clinical trial and re-apply this knowledge and skills in their future research. Taken together, the results in the paper imply that individual-level learning-by-doing is a crucial mechanism for enhancing clinical trial capacity in Sub-Saharan Africa.

Our results (Table 10) are remarkably robust to inclusion of alternative modifiers that could influence subsequent trial participation that also could be correlated with prior trial experience, such as applied nature of research, OECD connections, just to name a few.

[Table 10 about here.]

5 Conclusion

In this paper, we examine how a scientist's experience gained through participating in a research project spills over to follow on innovative projects. 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 and that it is not linked to repeated trials with the same sponsor. 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. Compared to the trial novices, repeat trialists are also significantly more likely to use pre-existing knowledge and skills in their subsequent research.

Taken together, our findings imply that the knowledge and skills developed in a first clinical trial can spill over to subsequent trials—and that this spillover can occur between different funding organization's projects. Since an initial funder may not internalize the full benefit of scientists' knowledge and skills development, this may lead to an under-supply of resources to these projects. This may be particularly impactful in resource poor environments where there are fewer funding opportunities to engage in an

initial clinical trial to develop knowledge and skills than in countries with more investment in scientific research. Policymakers who seek to increase the scientific and trial capacity in resource poor environments to enable the development of new drugs and vaccines may need to address the under-supply of resources through public investment in initial projects that facilitate learning-by-doing. Global cooperation between funding organizations may be needed to ensure that the value of spillovers can be better accounted for in funding decisions.

Interestingly, while our findings suggest that there may be intra-personal spillovers between funders' projects through individual scientists' learning, we find little evidence of wider inter-personal spillovers from scientists working on clinical trials to their colleagues. This suggests that policymakers need to invest in developing trial capacity by investing in individual learning—particularly through initial projects amongst less experienced researchers.

References

Pierre Azoulay, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat. Public r&d investments and private-sector patenting: evidence from nih funding rules. *The Review of economic studies*, 86(1):117–152, 2019.

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

Kevin A Bryan and Yasin Ozcan. The impact of open access mandates on invention. *Review of Economics and Statistics*, 103(5):954–967, 2021.

Eric Budish, Benjamin N Roin, and Heidi Williams. Do firms underinvest in long-term research? evidence from cancer clinical trials. *American Economic Review*, 105(7):2044–2085, 2015.

Joseph A DiMasi, Ronald W Hansen, Henry G Grabowski, and Louis Lasagna. Cost of innovation in the pharmaceutical industry. *Journal of health economics*, 10(2):107–142, 1991.

Jeffrey L Furman and Megan J MacGarvie. Academic science and the birth of industrial research laboratories in the us pharmaceutical industry. *Journal of Economic Behavior & Organization*, 63(4):756–776, 2007.

Jeffrey L Furman, Margaret Kyle, Iain M Cockburn, and Rebecca Henderson. Public & private spillovers, location and the productivity of pharmaceutical research, 2006.

Zvi Griliches. The search for r&d spillovers. *Scandinavian journal of economics*, 94(0):29–47, 1992.

Sabrina T Howell. Financing innovation: Evidence from r&d grants. *American economic review*, 107(4): 1136–1164, 2017.

Adam B Jaffe. Real effects of academic research. *The American economic review*, pages 957–970, 1989.

Michael Kremer and Rachel Glennerster. *Strong medicine: creating incentives for pharmaceutical research on neglected diseases*. Princeton University Press, 2004.

Margaret K Kyle and Anita M McGahan. Investments in pharmaceuticals before and after trips. *Review of Economics and Statistics*, 94(4):1157–1172, 2012.

Margaret K Kyle, David B Ridley, and Su Zhang. Strategic interaction among governments in the provision of a global public good. *Journal of Public Economics*, 156:185–199, 2017.

Danielle Li, Pierre Azoulay, and Bhaven N Sampat. The applied value of public investments in biomedical research. *Science*, 356(6333):78–81, 2017.

Thomas J Moore, Hanzhe Zhang, Gerard Anderson, and G Caleb Alexander. 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, 2018.

Enrico Moretti, Claudia Steinwender, and John Van Reenen. The intellectual spoils of war? defense r&d, productivity and international spillovers. *NBER Working Paper*, 26483, 2021.

Kyle Roy Myers and Lauren Lanahan. Research subsidy spillovers, two ways. In *Academy of Management Proceedings*, volume 2021, page 10397. Academy of Management Briarcliff Manor, NY 10510, 2021.

Richard R Nelson. The simple economics of basic scientific research. *Journal of political economy*, 67(3): 297–306, 1959.

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

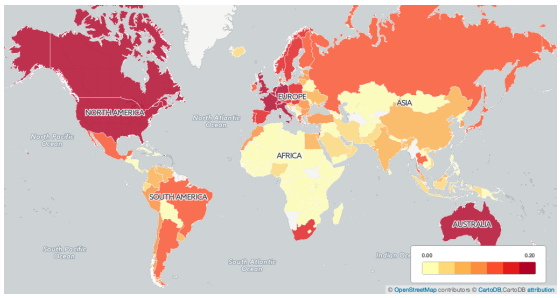
Fabio A Thiers, Anthony J Sinskey, Ernst R Berndt, et al. Trends in the globalization of clinical trials. *Nature Reviews Drug Discovery*, 7(1):13–14, 2008.

Lynne G Zucker and Michael R Darby. Individual action and the demand for institutions: Star scientists and institutional transformation. *American Behavioral Scientist*, 40(4):502–513, 1997.

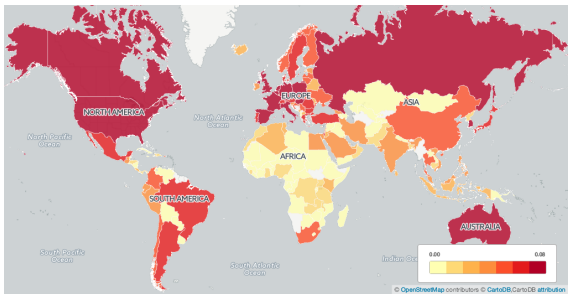
Lynne G Zucker, Michael R Darby, and Jeff S Armstrong. Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. *Management science*, 48(1):138–153, 2002.

Figure 1: Global locations of clinical trials over time

Panel A: 2000

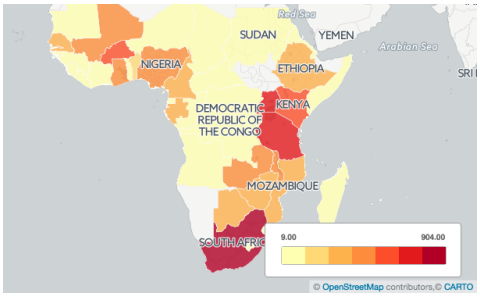


Panel B: 2015



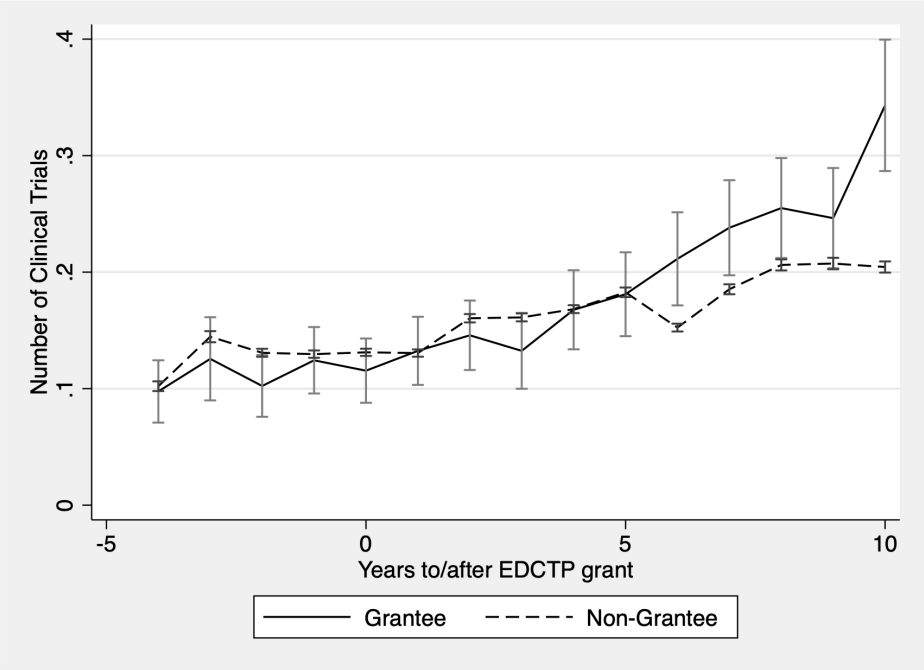
Notes: We plot the density of clinical trials indexed in Clinicaltrials.gov starting in 2000 (Panel A) and 2015 (Panel B) around the world.

Figure 2: Clinical trials in Sub-Saharan Africa



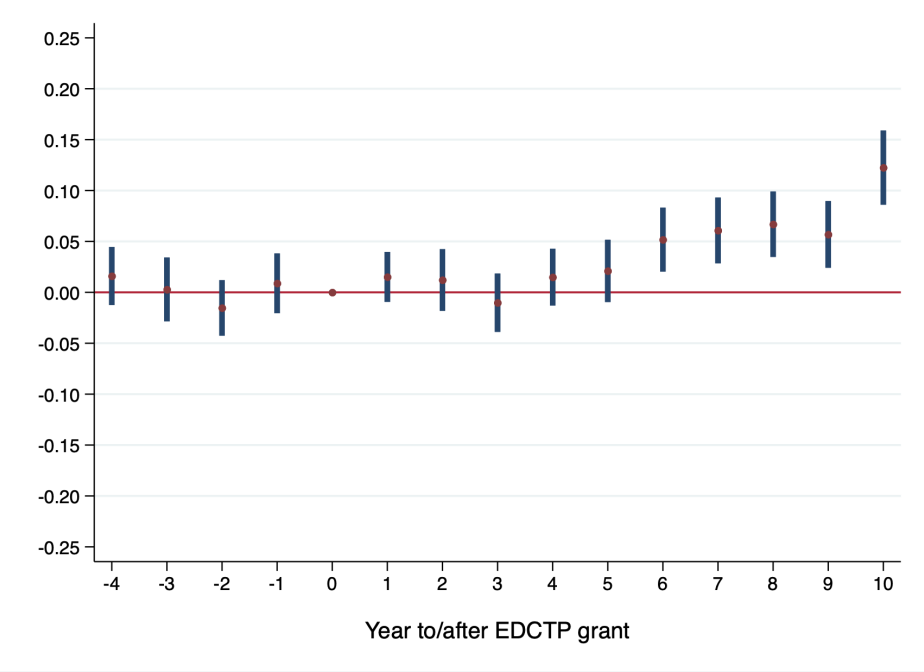
Notes: We plot the density of clinical trials indexed in Clinicaltrials.gov starting between 2000 and 2015 in countries in sub-Saharan Africa.

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



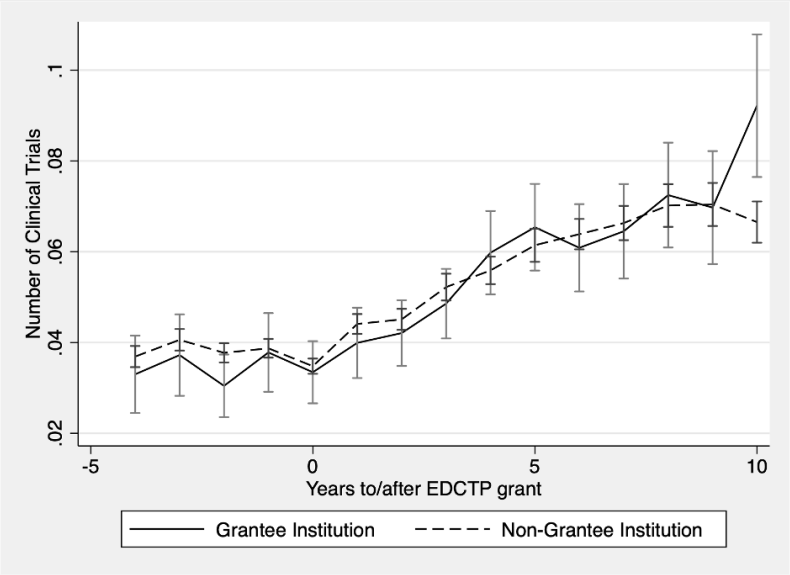
Notes: Raw trends of average outcomes per researcher for the are plotted for the five years before and and ten years after the grant.

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



Notes: Cefficient estimates stemming from conditional (scientist) fixed effects ordinary least squares specifications in which inverse hyperbolic sine outcomes 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. The 90% confidence interval robust standard errors clustered around the institution is plotted with solid bars.

Figure 5: Participation in clinical trials among non-EDCTP grantees in same institution as EDCTP grantee



Notes: Raw trends of average outcomes per researcher affiliated with the same institution as an EDCTP grantee are plotted for the five years before and ten years after the grant

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	0.323	0.358	-0.035	0.444
Number of Trials at Institution	41.849	43.652	-1.803	0.691
Size of Institution	227.80	248.95	-28.15	0.165
Number of OECD Co-authors	6.812	6.365	0.175	0.745
Number of Publications	6.288	5.596	0.692	0.006
Number of HIV/TB/Malaria Publications	2.239	1.423	0.816	0.000
Career Age in Treatment Year	10.361	10.009	0.351	0.274
Number of Scientists	823	144,475		

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

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

	Number of Clinical Trials		Number of non-EDCTP Clinical Trials	Number of Interventional Clinical Trials	Number of Phase 1 Clinical Trials	Number of Phase 2 Clinical Trials	Number of Phase 3 Clinical Trials	Number of Phase 4 Clinical Trials	Number of Observational Clinical Trials
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDCTP Grantee × Post-Grant	0.0371***	0.0346***	0.0372***	0.0238***	0.0126***	0.0039	0.0045	0.0129***	0.0121***
	(0.010)	(0.009)	(0.010)	(0.009)	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)
Total Observations	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637
Mean of Dep. Variable	0.0180	0.0180	0.0180	0.0141	0.0014	0.0023	0.0020	0.0020	0.0038
Author FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	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.
[b] Heteroskedastic robust standard errors are given in parentheses.

Table 3: Changes in clinical trial characteristics after receiving an EDCTP grant

	Industry Sponsored Clinical Trials	Non-Industry Sponsored Clinical Trials	EU Sponsored Clinical Trials	US Sponsored Clinical Trials	Africa Sponsored 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)	(7)	(8)	(9)
EDCTP Grantee × Post-Grant	-0.0023	0.0374***	0.0060*	0.0229***	0.0043	0.0236***	0.0172**	0.0269***	0.0053
	(0.003)	(0.009)	(0.003)	(0.007)	(0.003)	(0.005)	(0.008)	(0.008)	(0.004)
Total Observations	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637
Mean of Dep. Variable	0.0016	0.0164	0.0022	0.0075	0.0015	0.0018	0.0161	0.0067	0.0031
Author FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	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 4: Changes in scientists' research direction after receiving an EDCTP grant

	Number of Publications on Diseases	Number of Non-Disease Publications	Number of HIV, Malaria TB Publications	Number of Publications on Diseases Not HIV, Malaria TB	Number of Publications Trial Disease	Number of Publications Non-Trial HIV, Malaria or TB Disease	Number Basic Publications	Number Applied Publications
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EDCTP Grantee × Post-Grant	0.1216***	-0.0521***	0.1075***	-0.0174	0.1714***	-0.0395***	-0.0297**	0.0746***
	(0.013)	(0.014)	(0.019)	(0.023)	(0.018)	(0.012)	(0.015)	(0.014)
Total Observations	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637
Mean of Dep. Variable	0.3269	0.5491	0.2071	0.1198	0.0838	0.1233	0.3714	0.2893
Author FE	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	X	X	X	X	X
Annual Pubs	X	X	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 EDCTP grant impact by scientists' prior trial experience

	Number of Clinical Trials		Number of non-EDCTP Clinical Trials	Number of Interventional Clinical Trials	Number of Phase 1 Clinical Trials	Number of Phase 2 Clinical Trials	Number of Phase 3 Clinical Trials	Number of Phase 4 Clinical Trials	Number of Observational Clinical Trials
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDCTP Grantee × Post-Grant × Prior Trial	-0.0680**	-0.0681**	-0.0653*	-0.0805***	0.0061	-0.0307*	-0.0038	0.0148	0.0092
	(0.034)	(0.033)	(0.034)	(0.030)	(0.019)	(0.016)	(0.016)	(0.012)	(0.016)
EDCTP Grantee × Post-Grant	0.0526***	0.0499***	0.0521***	0.0420***	0.0114***	0.0108***	0.0057**	0.0098***	0.0100***
	(0.007)	(0.008)	(0.007)	(0.007)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Total Observations	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637
Mean of Dep. Variable	0.0180	0.0180	0.0180	0.0141	0.0014	0.0023	0.0020	0.0020	0.0038
Author FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	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. [b] Heteroskedastic robust standard errors are given in parentheses.

Table 6: Heterogeneity in EDCTP grant impact by scientists' career age

	Number of Clinical Trials		Number of non-EDCTP Clinical Trials	Number of Interventional Clinical Trials	Number of Phase 1 Clinical Trials	Number of Phase 2 Clinical Trials	Number of Phase 3 Clinical Trials	Number of Phase 4 Clinical Trials	Number of Observational Clinical Trials
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDCTP Grantee × Post-Grant × Career Age	-0.0023*	-0.0022*	-0.0024*	-0.0019*	-0.0004	-0.0005	-0.0003	0.0002	-0.0007
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
EDCTP Grantee × Post-Grant	0.0610***	0.0578***	0.0622***	0.0440***	0.0170***	0.0089*	0.0080*	0.0112***	0.0196***
	(0.013)	(0.013)	(0.013)	(0.012)	(0.006)	(0.005)	(0.004)	(0.004)	(0.005)
Total Observations	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637
Mean of Dep. Variable	0.0180	0.0180	0.0180	0.0141	0.0014	0.0023	0.0020	0.0020	0.0038
Author FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	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. [b] Heteroskedastic robust standard errors are given in parentheses.

Table 7: Heterogeneity in EDCTP grant impact by prior trial experience at the institution level

	Number of Clinical Trials		Number of non-EDCTP Clinical Trials	Number of Interventional Clinical Trials	Number of Phase 1 Clinical Trials	Number of Phase 2 Clinical Trials	Number of Phase 3 Clinical Trials	Number of Phase 4 Clinical Trials	Number of Observational Clinical Trials
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDCTP Grantee × Post-Grant									
× Prior Trial Institution	0.0101	0.0100	0.0115	0.0135	0.0027	-0.0071	0.0039	0.0048	0.0001
	(0.019)	(0.019)	(0.019)	(0.017)	(0.009)	(0.009)	(0.007)	(0.006)	(0.009)
EDCTP Grantee × Post-Grant	0.0340***	0.0315***	0.0336***	0.0196*	0.0117**	0.0061	0.0033	0.0114***	0.0120***
	(0.011)	(0.011)	(0.012)	(0.010)	(0.006)	(0.004)	(0.005)	(0.004)	(0.005)
Total Observations	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637	1832637
Mean of Dep. Variable	0.0180	0.0180	0.0180	0.0141	0.0014	0.0023	0.0020	0.0020	0.0038
Author FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X	X	X
Time Since Grant FE	X	X	X	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. [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	Number of OECD Coauthors	Number of African Coauthors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EDCTP Grantee × Post-Grant × Prior Trial	-0.0703	-0.0001	0.0878**	0.0867	-0.0234	0.1132	-0.0637
	(0.070)	(0.087)	(0.041)	(0.064)	(0.063)	(0.088)	(0.082)
EDCTP Grantee × Post-Grant	0.4396***	0.4686***	0.0514***	0.2332***	0.3377***	0.2583***	0.4976***
	(0.028)	(0.032)	(0.013)	(0.023)	(0.025)	(0.030)	(0.033)
Total Observations	1045087	1045087	1045087	1045087	1045087	1045087	1045087
Mean of Dep. Variable	0.9133	0.8260	0.1844	0.3219	0.5933	0.6836	0.9988
Author FE	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X
Career Age FE	X	X	X	X	X	X	X
Time Since Grant FE	X	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 9: Acquisition of knowledge and skills in clinical trials by prior trial experience

	Share of Trial Terms New-to-the-Scientist (in EDCTP Treatment Pubs)	Number of Repeated New-to-the-Scientist Trial Terms	Number of Repeated Not New-to-the-Scientist Trial Terms	Number of Repeated Not New-to-the-Scientist Trial Terms
	(1)	(2)	(3)	(4)
EDCTP Grantee x Prior Trial	-0.1115***	-0.0800	0.4955***	0.3862***
	(0.015)	(0.088)	(0.121)	(0.120)
Total Observations	971	6033	6033	6033
Mean of Dep. Variable	0.3478	1.0654	1.9097	1.9097
Year FE	X	X	X	X
Career Age FE	X	X	X	X
Time Since Grant FE	X	X	X	X
Terms Used Pre-EDCTP Trial				X

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

Notes: Sample in Model 1 only includes publications from trials linked to EDCTP treatment. Sample in other Models includes all observations with a publication among treated scientists with controls for number of publications in a given year. In Column 4, we additionally control for the number of terms that had been indexed to a scientist's papers before participating in the EDCTP trial. Heteroskedastic robust standard errors, clustered at the level of the scientist, are given in parentheses.

Table 10: Robustness

	Number of Clinical Trials					
	(1)	(2)	(3)	(4)	(5)	(6)
EDCTP Grantee × Post-Grant × Prior Trial	-0.0680**	-0.0634*	-0.0851**	-0.0648*	-0.0802**	-0.0686**
	(0.034)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)
EDCTP Grantee × Post-Grant	0.0526***	0.0789***	0.0442***	0.0486***	0.0458***	0.0529***
	(0.007)	(0.014)	(0.010)	(0.011)	(0.012)	(0.010)
EDCTP Grantee × Post-Grant × Career Age		-0.0026**				
		(0.001)				
EDCTP Grantee × Post-Grant × JIF Pubs			0.0013			
			(0.001)			
EDCTP Grantee × Post-Grant × OECD Pubs				0.0018		
				(0.003)		
EDCTP Grantee × Post-Grant × Applied Pubs					0.0023	
					(0.003)	
EDCTP Grantee × Post-Grant × HIV, Malaria, TB Pubs						-0.0001
						(0.004)
Total Observations	1832637	1832637	1832637	1832637	1832637	1832637
Mean of Dep. Variable	0.0180	0.0180	0.0180	0.0180	0.0180	0.0180
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.