An Overview of Field Experiments in Health in Developing Countries

Jorge Yépez[1]

[1] *Instituto de Altos Estudios Nacionales*

Correspondence concerning this article should be addressed to Jorge Yépez Instituto de Altos Estudios Nacionales (Ecuador, Quito, Ecuador, Av. Amazonas N37-271), e-mail: yepezeta@gmail.com

Medical science recognizes that patients do not always respond to treatment in a way physicians expect; moreover, different people respond to treatment in different and sometimes unpredictable ways. These differences can be attributed to observable patients' characteristics, such as age and gender, but also unobservable characteristics can influence the results, such as lifestyle or an unreported health condition. To overcome this problem, during the twentieth century, medical researchers and statisticians began extensively using randomized controlled trial (RCT) methods to assess the effect of the treatment and to separate this effect from other patients' characteristics that could affect the results. During the last decades, social scientists, willing to go beyond correlational analysis to provide insights on causation, have adopted a similar approach. This paper gives a short overview of RCT in development economics with a special attention to programs related to health in developing countries.

Keywords: randomized control trial, field experiments, impact evaluation, development economics, health

Introduction

Let me take you back in time, to the period of the Old Testament when Daniel, being a vegetarian, refuses to eat the royal delicacies offered by the King Nebuchadnezzar. After hearing this, the king settles that for ten days Daniel and three young men will eat only vegetables, and will then be compared with others who have eaten royal delicacies. At the end of the experiment, Daniel and the other three men were healthier than the comparison group subjects, so they were allowed to remain vegetarian. Although this can be considered the first treatment vs control group study, we have to go forward, to the year 1747, when a ship's surgeon named James Lind added randomization to his experiment. The surgeon decided to compare different scurvy treatments. Lind randomized six treatments, each one was given to a pair of sailors. The first group got cider, the second sulphuric acid, the third vinegar, the fourth seawater, the fifth a mixture of nutmeg, garlic, mustard seeds, horseradish, balsam of Peru and gum myrrh, and the sixth group got two oranges and one lemon. All the participants were fed the same diet and were kept in the same part of the ship (Leigh, 2018, p. 3). In 1753, Lind published his results noting that "the most sudden and visible good effects were alleged from the use of the oranges and lemons" (Meldrum, 2000).

Main Body

The expansion of scientific methods in the nineteenth century encouraged the application of experimental research. Pierre Louis stressed the importance of selecting control subjects as similar as you could find to the treatment subjects. Possibly Johannes Fibiger's 1898 serum treatment study was the first documented randomized controlled trial (RCT) experiment. Diphtheria patients admitted to the Copenhagen hospital were allocated into the treatment group where subjects received serum injections, while the control group subjects received traditional treatment.

By the beginning of the twentieth century, the profusion of new medicines, coupled with manufacturers' claims that their products were the best in the market, motivated the creation of the Council on Pharmacy and Chemistry of the American Medical Association. The idea was to provide expert valuation on the drugs available on the market by the use of RCTs (Meldrum, 2000). Nowadays, the approval of a new drug requires an initial safety testing to be done on animals before being tested on people. Human testing is a three phase RCT process. At phase I, the drug's efficacy is tested on fewer than a hundred people. During phase II the drug is tested on a few hundred people. Finally, at phase III trials test

effectiveness on several thousand and compare the results with other drugs (Leigh, 2018, p. 28).

The use of RCTs in the field of development economics, assessing the impact of social programs, started in the 70s. The Ford Foundation and several federal government agencies created in 1974 the Manpower Demonstration Research Corporation (MDRC) to advance the understanding about what worked in social policy. Judith Gueron, the first MDRC research director, specified in her book Fighting *for Reliable Evidence* that the motivation for using RCT is to answer the following question: "is it possible to accurately and reliably measure the impact of a program in order to determine whether it is worth the money spent on it?" (Ogden, 2017, p. xx).

Although, at first only wealthy governments had the capability and budget to evaluate social policies using RCT, middle income countries started to follow suit. The Mexican government ran a very large scale RCT to evaluate a new conditional cash transfer program called PROGRESA in the early 90s. Around the same time, the 2019 Nobel laureate in economics Michael Kremer convinced a Duch NGO to fund a RCT to evaluate effect of textbooks in Kenyan primary school students' test scores (Glewwe et al., 1998). As Ogden (2017, xxi) stated in his book Experimental Conversations Kremer's innovation was not just in bringing RCTs into a new environment but seeing that while it was true that running large-scale experiments with government funding was usually impossible in less developed countries, there was another path to conducting such trials: working with NGOs (Ogden, 2017, p. xxi).

Roughly 30 years later, thousands of RCTs have been conducted in developing countries. Kremer later worked with Abhijit Banerjee, Esther Duflo and Sedhil Mullainathan to create the Jameel-Poverty Action Lab (J-PAL) at MIT, an ONG focused on applying the findings of RCTs to policy and program design. As the other 2019 Nobel laureates Banerjee and Duflo (2012, 8) write, "the cleanest way to answer such questions [of causality] is to mimic the randomized trials that are used in medicine to evaluate the effectiveness of new drugs".

Nowadays, RCT is considered the gold standard of impact evaluation. The method uses a process of selecting beneficiaries of a program at random, such as flipping a coin to decide who receives and who does not receive the benefit delivered by a social program. Under randomized assignment, every eligible unit has the same probability of receiving the benefit. As long as the number of randomized units is sufficiently large, the process will produce groups

that have statistically equivalent averages for all their characteristics. In other words, the treatment and comparison groups are statistically identical. The researcher can test this assumption using statistical tests such as t-tests for difference in means. For example, observed characteristics, such as sex or the educational level, should be on average the same for both groups. If this is true, the researcher can assume that characteristics that are more difficult to observe (unobserved variables), such as motivation or preferences, are also the same for both the treatment and comparison groups (Gertler et al., 2016).

Thus, after randomization, treatment and comparison groups will be identical in every possible way (observed and unobserved characteristics) except for the fact that only the treatment group will receive the program. After the program is launched, if we observe differences in outcomes between the treatment and control groups, those differences can be explained only by the program. The impact of a program could simply be estimated by taking the difference between the mean outcome of the treatment group and the mean outcome of the comparison group. However, in order to have more precise estimations – smaller standard errors, researchers usually evaluate the impact using a multivariate regression model (Gertler et al., 2016).

Scientist in the field of development economics applied RCTs in different areas, such as: agriculture, crime, violence, conflict, education, environment, finance, labor markets and health (J-PAL, n.d.). To overview a few examples in the health development area, I will start with a study of Michael Kreamer and Edward Miguel. The authors evaluated a deworming program where schools were treated with worm medication. Using a randomize evaluation across a sample of 75 schools with a total enrollment of more than 30,000 students between the ages of 6 and 18, in southern Busia, a poor rural region in western Kenya, the authors found that treated schools exhibited improved health compared to non-treated schools. The idea of treating schools, instead of households, rely on the fact that after receiving the medication children worm load will decrease. In addition, children that were not exposed to the program but lived nearby treated children would also no longer come in contact with the treated children's worms. The authors found substantial spillover effects, where untreated children in schools located close to the treatment schools show positive effects compared to students from farther away schools (Miguel and Kremer 2004).

(Kremer et al., 2011) also evaluated a simple program in Kenya that aimed to reduce contaminated water and diarrheal incidence by sealing off water spring sources

and encasing them in concrete. Spring protection helps water streams directly into a bucket rather than the ground or people's hands where it is vulnerable to contamination. Researchers randomly selected 100 springs for the treatment group where the water sources were encased in concrete, and 100 springs for the control group. The researchers found that the simple spring protection reduced the presence of E. coli by 66 percent compared to untreated springs. Moreover, diarrheal incidence for children under three years old fell by one-quarter relative to the control group.

During the last year, recognizing the role of psychological factors in decision making, experimental researches tended to include the use of incentives to boost the effect of a program. For instance, (Banerjee et al., 2010) evaluated a program in Udaipur, rural India, that improved children immunization rate.

In India parents could face a lot of problems when trying to immunize their children, even when immunization is free. For instance, the average household is quite far from the nearest clinic, so reaching the clinic can be time-consuming. In addition, clinics are usually closed due to the fact that nurses often skip work. Moreover, parents could not fully understand the benefits of immunization and tended to see the process as inconvenient. Considering all the above, this program was created to ease the immunization process for the parents. The program set up immunization camps where the presence of a nurse and an assistant at the camp was documented by a photo with a time and date stamp. In addition, special incentives were offered to the parents. A 1 kg. bag of dried beans was offered per immunization, and a set of plates were given to the parents after their child was fully immunized. The authors randomly assigned 30 villages to the treatment group where well-publicized immunization camps where held, 30 other villages were assigned to a second treatment group where similar camps were held plus parents were also offered the incentives to immunize their children, and 74 villages served as a control group. The study found out that full immunization rates increased from 6 percent in control group to 18 percent in villages with camps but no incentives and 39 percent in villages with camps and incentives.

Similar to India, Uganda health workers also tend to be absent. To improve accountability, (Björkman & Svensson, 2009) evaluated the impact of a community-based monitoring program. Fifty public health dispensaries in nine districts covering all four regions in Uganda were randomly assigned either to a program or to a comparison group. In the program communities, researchers generated report cards for each dispensary that were disseminated in the communities. The cards included information on the dispensary's service quality relative to neighboring facilities. The program also brought community members and health workers together to create a shared action plan on what had to be done to improve the service. The program led to significant improvements in health outcomes after one year, such as a 33 percent reduction on under-five mortality rate and an increase on the average weight-for-age of infants. The program also reported a 10 percent improvement in attendance as well as a 9 percent reduction in waiting time and 28 percent reduction in absenteeism.

Conclusion

To conclude, in the health sector, development economics uses RCTs to evaluate programs and policies effectiveness improving the quality and the reach of health services. Although the methodology was first used in the medical field, it is currently extended to different areas such as the evaluation of public policies. During the last years, with the recognition of the role of behavioral economics, field experiments tend to include the incentives or nudges to increase the take-up of a program.

References

Banerjee, A. V., & Duflo, E. (2012). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. Public Affairs.

Banerjee, A. V., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural India: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ (Clinical Research Ed.)*, *340*, c2220. https://doi.org/10.1136/bmj.c2220

Björkman, M., & Svensson, J. (2009). Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda. *The Quarterly Journal of Economics*, *124*(2), 735–769.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice, Second Edition*. World Bank. https://openknowledge.worldbank.org/handle/10986/25030

Glewwe, P., Kremer, M., & Moulin, S. (1998). Textbooks and test scores: Evidence from a prospective evaluation in kenya. In *Natural Field Experiments* (Nº. 00255; Natural Field Experiments). The Field

Experiments Website. https://ideas.repec.org/p/feb/natura/00255.html

J-PAL. (n.d.). *Policy Insights*. The Abdul Latif Jameel Poverty Action Lab (J-PAL). Retrieved January 11, 2022, from https://www.povertyactionlab.org/policy-insights

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions. *The Quarterly Journal of Economics*, *126*(1), 145–205. https://doi.org/10.1093/qje/qjq010

Leigh, A. (2018). *Randomistas: How Radical Researchers Are Changing Our World*. Yale University Press.

Meldrum, M. L. (2000). A brief history of the randomized controlled trial: From oranges and lemons to the gold standard. *Hematology/Oncology Clinics of North America*, *14*(4), 745–760. https://doi.org/10.1016/S0889-8588(05)70309-9

Miguel, E., & Kremer, M. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, *72*(1), 159–217. https://doi.org/10.1111/j.1468-0262.2004.00481.x

Ogden, T. N. (Ed.). (2017). *Experimental Conversations: Perspectives on Randomized Trials in Development Economics*. MIT Press.

Обзор полевых экспериментов в области здравоохранения в развивающихся странах

Йепес Хорхе[1]

[1] Институт перспективных национальных исследований

Корреспонденция, касающаяся этой статьи, должна быть адресована Йепес Х. Институт перспективных национальных исследований, адрес: Кито, Эквадор, авеню Амазонас N37-271, e-mail: yepezeta@gmail.com

Медицинская наука признает, что пациенты не всегда реагируют на лечение так, как ожидают врачи; более того, разные люди реагируют на лечение по-разному и иногда непредсказуемо. Эти различия могут быть связаны с параметрами, которые поддаются наблюдению, такими как возраст и пол пациентов, но на результаты могут влиять и параметры, которые не поддаются наблюдению, такие как образ жизни пациента, или отсутствие данных о состоянии его здоровья. Чтобы преодолеть эту проблему, в течение двадцатого века медицинские исследователи и статистики начали широко использовать методы рандомизированных контролируемых испытаний (РКИ) для оценки эффекта лечения и выделение полученного результата от других предпосылок развития болезни, которые могли повлиять на эффект от лечения. В течение последних десятилетий социологи, желающие выйти за рамки корреляционного анализа, чтобы получить представление о причинно-следственных связях, приняли аналогичный подход. В этой статье дается краткий обзор РКИ в экономике развивающихся стран с особым вниманием к программам, связанным со здоровьем.

Ключевые слова: рандомизированное контрольное испытание, полевые эксперименты, оценка воздействия, экономика развивающихся стран, здоровье

Литература

Banerjee, A. V., & Duflo, E. (2012). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. Public Affairs.

Banerjee, A. V., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural India: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ (Clinical Research Ed.)*, *340*, c2220. https://doi.org/10.1136/bmj.c2220

Björkman, M., & Svensson, J. (2009). Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda. *The Quarterly Journal of Economics*, *124*(2), 735–769.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice, Second Edition*. World Bank. https://openknowledge.worldbank.org/handle/10986/25030

Glewwe, P., Kremer, M., & Moulin, S. (1998). Textbooks and test scores: Evidence from a prospective evaluation in kenya. In *Natural Field Experiments* (№. 00255; Natural Field Experiments). The Field Experiments Website. https://ideas.repec.org/p/feb/natura/00255.html

J-PAL. (n.d.). *Policy Insights*. The Abdul Latif Jameel Poverty Action Lab (J-PAL). Retrieved January 11, 2022, from https://www.povertyactionlab.org/policy-insights

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions. *The Quarterly Journal of Economics*, *126*(1), 145–205. https://doi.org/10.1093/qje/qjq010

Leigh, A. (2018). *Randomistas: How Radical Researchers Are Changing Our World*. Yale University Press.

Meldrum, M. L. (2000). a brief history of the randomized controlled trial: From oranges and lemons to the gold standard. *Hematology/Oncology Clinics of North America*, *14*(4), 745–760. https://doi.org/10.1016/S0889-8588(05)70309-9

Miguel, E., & Kremer, M. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, *72*(1), 159–217. https://doi.org/10.1111/j.1468-0262.2004.00481.x

Ogden, T. N. (Ed.). (2017). *Experimental Conversations: Perspectives on Randomized Trials in Development Economics*. MIT Press.