

The fog of development: evaluating the Millennium Villages Project



Shira Mitchell and colleagues¹ endline evaluation of the Millennium Villages Project (MVP) in *The Lancet Global Health* marks an important chapter in our understanding of Africa's meandering path towards health and economic development. Originally conceived to show the power of an integrated multisector approach to ending poverty and its associated ills, the project had its share of heated debates.^{2,3} The centrally planned approach that included provision of a streamlined basket of goods to each village was said to promote solutions derived from aloof economic models insensitive to local customs and constraints.⁴ Furthermore, the absence of upfront plans for evaluating the project hinted at overconfidence in the righteousness of the approach, which was inconsistent with the prevailing equipoise about the effectiveness of approaches to improving the health and wealth of the world's poorest. Over the past decade, many important approaches to development, including trials of multisector aid programmes, have been subjected to controlled trials wherein village clusters are randomly assigned to intervention and control groups.⁵ The MVP, however, selected villages deliberately and was set up without routine monitoring of comparison villages.

For that reason, the analysis by Mitchell and colleagues represents the culmination of heroic efforts to identify the treatment effect of the MVP against stacked methodological odds. The selection of comparison villages, instead of being based on a coin toss or a baseline evaluation, involved digging back and matching a host of key features circa 2005. The information for the comparison features was derived from satellite data and household surveys, among other sources, but did not include most of the endline outcomes. We are, therefore, unaware of baseline differences in the outcomes between the MVP and comparison villages. Matched villages were then surveyed for the same outcomes as the MVP villages in 2015. Although endline-only comparison is commonly acceptable for randomised trials (because randomisation presumably balances baseline observables and unobservables), this approach is tenuous in a retrospective study, especially because MVP villages were not selected at random.

The non-random selection means that, even if baseline observables are well matched, evolving differences might be unrelated to the intervention. Suppose a village in Kenya was selected for MVP participation because of its highly engaged local government in 2005. At the time, the coverage of bednets was similarly low in the MVP and in the (eventual) comparison village, because few bednets were available. 10 years on, the MVP village might surge ahead of the comparison village on bednet coverage and child mortality, not because of the MVP interventions, but because of its engaged and responsive local government.

Nevertheless, the publication of this study is an important bookend to a decade of rhetoric about the effectiveness of the MVP. The numbers—or at least the indices—are in. On 30 of the 40 measures, the MVP villages are better off, on average, than the comparison villages. This finding is reassuring given the amount of money that has been invested in building fundamental capacities in agriculture, education, infrastructure, and health (about US\$600 million in direct investments, amounting to \$120 per person per year over 10 years for 500 000 people living in the Millennium villages; author's estimation). The one area in which unequivocal and substantial benefits are observed is maternal health, including contraception use, antenatal care, and use of skilled birth attendants. The effect sizes for other key outcomes, including child malnutrition and mortality, are small, heterogeneous, and unstable (the overall child mortality effect seems driven by the comparison in poorly matched Nigeria). Moreover, although not the focus of health analyses, the end of poverty—arguably the *raison d'être* of the entire project—was no closer in the MVP villages than in the comparison villages.

An important point of context is that the MVP operated during a decade of remarkable improvements throughout the continent, making this evaluation a horse race against serious contenders. Similar to the joke about the camper who goes to sleep with his boots on to avoid being eaten by a lion (not because he'll be able to outrun the lion, but because he'll be able to outrun his tent-mates), this endline evaluation assesses the extent to which the MVP villages managed to outrun

See [Articles](#) page e500

comparison villages during a decade when child mortality halved, economic growth was robust, and many of the technologies promoted by the MVP diffused widely. The success of the MVP might have been partly undone by its next of kin millennial: the Millennium Development Goals (MDGs). The MDGs galvanised an enormous amount of goodwill and resources towards many of the same goals as the MVP. Unlike the MVP, however, the MDGs operated on a global scale, affecting national and international institutions. Ministries of health, local and global non-governmental organisations, and global health institutions organised, strategised, and monitored progress around the MDGs. If the equivocal success of the MVP is a reflection of the runaway success of the MDGs, so much the better for the world's poor, even if the formula of the MVP is not the magic bullet to achieving sustainable development. Their failure to become islands of progress is, nevertheless, a valuable lesson.

Sometimes in the course of human events it becomes necessary to subject important programmes and ideas to a rigorous evaluation.⁶ The authors deserve to be

congratulated for squeezing as much juice as possible from this academic third rail, and for transparently and courageously telling the world: this was it, this is the best we can say about the MVP.

Eran Bendavid

Stanford University, Stanford, CA 94305, USA

ebd@stanford.edu

I declare no competing interests.

Copyright © The Author(s). Published by Elsevier Ltd. This is an open access article under the CC-BY-NC-ND 4.0 license.

- 1 Mitchell S, Gelman A, Ross R, et al. The Millennium Villages Project: a retrospective, observational, endline evaluation. *Lancet Glob Health* 2018; **6**: e500–13.
- 2 Pronyk P. Errors in a paper on the Millennium Villages project. *Lancet* 2012; **379**: 1946.
- 3 Bump JB, Clemens MA, Demombynes G, Haddad L. Concerns about the Millennium Villages project report. *Lancet* 2012; **379**: 1945.
- 4 Munk N. The idealist: Jeffrey Sachs and the quest to end poverty. New York, NY: Doubleday, 2013.
- 5 Banerjee A, Duflo E, Goldberg N, et al. A multifaceted program causes lasting progress for the very poor: evidence from six countries. *Science* 2015; **348**: 1260799.
- 6 Clemens MA, Demombynes G. When does rigorous impact evaluation make a difference? The case of the Millennium Villages. *J Dev Effect* 2011; **3**: 305–39.