A conversation with Nathan Lo, April 4, 2019

Participants

- Nathan Lo, PhD – MD-PhD program, Stanford University School of Medicine
- Amar Radia – Senior Research Analyst, GiveWell

Note: These notes were compiled by GiveWell and give an overview of the major points made by Dr. Lo and related discussion.

Summary

GiveWell spoke with Dr. Lo of Stanford University School of Medicine as part of a review of its process for adjusting worm intensity when calculating the cost-effectiveness of charities' deworming programs.[1] Conversation topics included the conversion of worm prevalence into worm intensity, the advantage of utilizing data on the proportions of children with light, moderate, and heavy intensity over mean intensity, and dynamic modeling of worm burden averted through deworming.

Conversion of worm prevalence into worm intensity

In the appendix of Lo et al. 2016, Dr. Lo and his colleagues present a model for converting worm prevalence into worm intensity. However, out-of-sample testing of this model performed by GiveWell using prevalence data from some of its top charities yielded systematic underestimation of intensity (actual intensity values were available for comparison). Potential explanations for these results include:

- **Differences in prior treatment** – Data used to create Lo et al. 2016's predictive model was gathered from communities where no mass drug administration (MDA) of deworming tablets had occurred. If baseline data used by GiveWell originates from settings in which MDA had occurred and a subset of the population had repeatedly missed treatment (a common issue), worm burden would have been concentrated in this subset—which Dr. Lo believes might cause the model to underestimate intensity in these settings. On the other hand, deworming treatment might be expected to reduce mean intensity disproportionately more than prevalence of any infection. This would go in the other direction, leading the model to overestimate intensity in areas with prior deworming.
- **Uncertainty intervals** – Lo et al. 2016's model for converting prevalence into intensity provides uncertainty intervals for intensity estimates, which account for heterogeneity between settings. Although the model tends to underestimate intensity when applied to GiveWell's prevalence data, the resulting intensity estimates mostly still fell within uncertainty intervals.

[1] E.g. here.

- **Limited sample size** – Parameters for the predictive model were created using data from a search of previous studies. As further data has become available, the precision of the model could potentially be improved.

Advantage of utilizing data on the proportions of children with light, moderate, and heavy intensity over mean intensity

GiveWell currently emphasizes mean intensity of worm infection—rather than proportions of light, moderate, and heavy intensity infections—in its approach to calculating the cost-effectiveness of deworming programs. However, Dr. Lo suggested that proportions of light, moderate, and heavy intensity infections could be stronger indicators of the potential effect of deworming on human health and development. GiveWell should therefore consider focusing on indicators such as the proportion of children with a moderate or heavy worm infection.

Dr. Lo believes that it is the presence of moderate or heavy worm infections during childhood is a crude, but useful distinction for disease morbidity. Very light infections are likely to have more limited effects relative to higher intensity infection, and the marginal effects of worm intensity are likely to diminish beyond a certain point. Furthermore, the duration of exposure is also relevant.

Using mean intensity in the community as the key indicator of morbidity could therefore be misleading. For example, a high mean intensity of infection in a group of 10 children could indicate high overall morbidity. However, it could be that eight of the children have light intensity infections or no infection and two of the children have extremely heavy intensity infections, in which case morbidity effects should be expected in only 20% of the population. Therefore, accounting for the dispersion is important and reporting prevalence of low, moderate, and heavy infection intensity is more reliable.

The egg thresholds used to distinguish between light, moderate, and heavy intensity are relatively arbitrary. However, matching proportions of a population to categories of worm intensity is still more informative than simply looking at mean intensity across the population.

Dynamic modeling of worm burden averted through deworming

Dr. Lo believes that the strongest method for projecting worm burden averted through deworming would be a dynamic model that could incorporate factors such as length of treatment, treatment coverage, interruptions in treatment, time elapsed since baseline data was recorded, and counterfactual results. GiveWell foresees three main use cases for this type of model:

- **Use case 1** – For use case 1, GiveWell would have baseline data on intensity and prevalence and would be considering funding a program that plans to treat an area with no prior history of MDA. The model for this case would be calibrated to underlying (baseline) worm intensity in that geography, expected coverage, reinfection rates for individual species of worms, and other variables and would then project forward the

benefits of the program on worm burden. Counterfactual results would likely be equivalent to baseline data and would therefore be simple to incorporate.

- **Use case 2** – For use case 2 (the most typical case), GiveWell would have baseline data on intensity and prevalence and would be considering providing additional funding to a program that has been operating for some amount of time. The model for this case would be largely similar to the model for use case 1, although further calibration would be required to incorporate the impacts of previous treatment and calculation of counterfactual results would be more complex.
- **Use case 3** – For use case 3, GiveWell would not have baseline data on intensity and prevalence and would be considering providing additional funding to a program that has been operating for some amount of time. The model for this case would be similar to use case 2, although calculation of counterfactual results would require using available data on treatment history (e.g. frequency, coverage, duration, intensity during treatment) to estimate baseline intensity and then project forward. Alternatively, baseline intensity could be calculated by collating data from other settings that have been exposed to similar amounts of treatment.

Opportunities for using dynamic models

Using dynamic models for different cases relevant to funding a deworming program would require an individual with familiarity with the models and programming skills. Options for taking this kind of work forward include: (i) GiveWell contracting someone with relevant skills and knowledge; (ii) GiveWell developing such skills and knowledge in-house; (iii) the development of a tool with a clear user interface that would be easy for GiveWell staff to use.

One individual that could be beneficial for GiveWell to speak with is Dr. Deirdre Hollingsworth, who leads the NTD Modelling Consortium and has been involved in the creation of user-friendly models.

All GiveWell conversations are available at
http://www.givewell.org/research/conversations