Reference

A critical review of GiveWell's 2022 cost-effectiveness model — EA Forum

Indirect malaria deaths

"There is a fairly serious error in the Original model where GiveWell assume that a malaria death prevented by Malaria Consortium also prevents 0.5 indirect deaths… and then forgets to apply this to AMF."

After looking into this issue, we do not believe that Froolow's claim is correct.

In the version of the CEA Froolow looked at last year, we calculated the impact of SMC on mortality by estimating the impact of SMC on malaria incidence, translating our incidence estimate into an estimate of the effect of SMC on malaria mortality, then accounting for indirect malaria deaths. For AMF, we instead rely on a meta-analysis result estimating the impact of insecticide-treated bed nets on all-cause mortality (see here), which we then adjust to account for changes in the proportion of all-cause mortality due to malaria during the trials compared to areas where AMF works today. Because we started with an all-cause mortality effect (rather than a malaria-specific one), we don't need to make an additional adjustment for indirect malaria deaths.

As a side note, we have since updated our method for estimating malaria mortality for separate reasons to a method more similar to our SMC method (translating reductions in malaria incidence to reductions in malaria mortality, including estimates of indirect malaria deaths). See version 3 of our 2023 CEA here.

Assumption about resource sharing differs between cash transfers and deworming

"The assumption of how income is aggregated across the household is different for different charities. In Cash Transfer charities it is assumed that the money from the cash transfer is divided across all 4.7 members of the household, whereas in the Deworming charities it is assumed that the money generated by the intervention is multiplied by two to account for resource-sharing within a household. This seemed very inconsistent to me, so I normalised everything to Cash Transfer approach. This has a big impact on any intervention which generates consumption income."

After looking into this issue, we do not believe that Froolow's claim is correct.

For a lump sum cash transfer program (GiveDirectly), we assume that the cash transfer results in equal increases in consumption across all household members. We divide the transfer amount ($1,000) by the average household size (4.7) to get a transfer size per person of ~$212. We then estimate increases in consumption per person based on the amount of the transfer that is consumed vs. invested and the expected return on that investment (see here).

For deworming, we intentionally use a different method for estimating per capita increases in consumption to account for the fact that there may be multiple wage-earners in a household. For example, in a household with one wage-earner, a 10% increase in consumption could allow all individuals to consume 10% more. However, in a two-wage earner household, a single individual earning 10% more would not be able to raise all individual's consumption by 10% (rather all individuals could have their consumption raised by a smaller amount).

This is described in the cell note here:

> "If a person treated for worms earns additional income and supports a family, then multiple people may benefit—not just the person who was dewormed.
>
> In a multi-person household with one wage earner, a 10% increase in wages could enable every member of the household to consume 10% more. However, many households will have multiple wage earners, and household size may change over time.
>
> A rough model for estimating a value for this parameter is available at https://docs.google.com/spreadsheets/d/112uuyYt6QLRZuJojwz6fHv4JQ-GHNeIpiT-SauY3kmM/edit#gid=0. The appropriate value for this parameter will depend on many uncertain factors (e.g. household composition and how household composition changes overtime). We currently use a default value of 2.0. Our rough model suggests values close to 2 under a range of reasonable assumptions."

Leverage and funging adjustment for GiveDirectly

"GiveDirectly has no leverage / funging adjustment, while every other charity does"

We don't believe it would be appropriate to include a leverage and funging adjustment for GiveDirectly.

For Malaria Consortium, for example, we think it's appropriate to include a leverage and funging adjustment because we think there's a meaningful chance that, in absence of GiveWell, other actors like the Global Fund or PMI might contribute more to SMC ("funging"). We also think that there's a chance that GiveWell's support of Malaria Consortium leads domestic governments to spend more on SMC than they otherwise would have ("leverage").

Our understanding is that these types of issues seem unlikely to apply to GiveDirectly's core program. We don't think that governments are contributing part of the costs of implementing GiveDirectly's program (so we don't think that GiveDirectly is leveraging government funds), and we don't think that there's a meaningful chance that other funders or governments would contribute more to cash transfer programs in absence of GiveWell.

Spillover effects for cash transfers for New Incentives

"When cash is transferred by GiveDirectly there is a 5% reduction in the amount transferred due to 'negative spillover effect'. However, when cash is transferred by New Incentives this deflator is not applied. People are assumed to invest a certain percentage of money transferred by GiveDirectly, but there is no functionality to invest money transferred by New Incentives."

We haven't looked into this issue in much depth, but our best guess is that it would not be appropriate to include adjustments for negative spillovers or investments for New Incentives.

New Incentives offers much smaller cash transfers than GiveDirectly: around ~$11 across all vaccine visits (or ~$8 if subtracting out transportation costs), compared to roughly ~$1000 per household for GiveDirectly. Based on the size of New Incentives' transfers, we'd guess that investment and potential negative spillovers would be minimal.

Discount rate for YLLs

"For example, instead of using the GiveWell assumptions about the value of life years lost, the model offers the possibility of using values derived from Lopez et al. These represent the present discounted value of life years lost for an intervention which saves a life at 5 years old, 10 years old, 15 years old and so on. The issue is that the discount rate is also an important input in the economic model – Lopez et al assume it is 3% while GiveWell assume it is 4% (calculations for this are, naturally enough, on a separate sheet). So GiveWell accidentally strongly commits to a philosophical position that the discount rate on life years is less than the discount rate for money."

We don't think this criticism applies to our recommended "GiveWell 2020" moral weights.

Froolow's criticism is related to the "conventional" moral weights adapted from Lopez et al. 2006 (see here). Our understanding is that those estimates do rely on a 3% discount rate for averted YLLs.

However, we only present the Lopez et al. moral weights for illustration—our recommended moral weights are in the "GiveWell 2020" column, which we describe in this document.

New Incentives uses a different discount rate

"New Incentives has a different discount rate for future life years than every other charity for reasons I can't understand."

It is an intentional choice to use different discount rates for increases in consumption and future averted deaths.

In our CEAs, the discount rate of 4% is only applied to future increases in consumption (e.g., future gains in income for deworming). For our top charities working in malaria prevention and vitamin A supplementation, we don't use a discount rate for averting future deaths because we think that the deaths averted by malaria nets, SMC, and vitamin A supplementation occur in the near-term.

For New Incentives, we also want to incorporate the benefits that infants who are vaccinated will receive when they are adults. For example, if an infant receives a BCG vaccine through New Incentives, which they wouldn't have received counterfactually, we think they have a lower chance of mortality from tuberculosis as adults than they would have in absence of New Incentives. Those future deaths are the ones we're applying discount rates to.

We don't think that the same reasoning for discounting future gains in consumption (discussed here) necessarily applies to discounting future deaths averted, so we intentionally use a different discount rate. We discuss our reasoning for using a 0.5% discount rate for future averted deaths here.

Modeling reductions in SMC effectiveness in areas with bed net distributions

"AMF loses a certain amount of effectiveness in some countries where Malaria Consortium operates due to some of their bednet distribution inevitably going to the same people as Malaria Consortium have targeted with Seasonal Malaria Chemoprevention. However, Malaria Consortium doesn't have a corresponding figure for effectiveness lost due to bednet distribution."

We don't believe that our approach is an error (though we have some uncertainty about our methodology for addressing this issue).

Our understanding from conversations with the Institute For Health Metrics and Evaluation (IHME), which produces the Global Burden of Disease (GBD) estimates we use is that their models of malaria prevalence and mortality (which we use in our malaria CEAs) take estimates of bed net coverage into account, but that they do not take estimates of SMC coverage into

account. Because SMC is not already accounted for in IHME's baseline malaria burden estimates, we include a separate adjustment.

IHME told us it may account for SMC coverage in future versions of the GBD—we're unsure about how well our estimates of the impact of SMC on malaria indicators will track IHME's estimates when SMC is introduced into the GBD model.

Double counting of some charity-level adjustments

"The adjustment for charity-level factors double-counts these issues. For example, GiveWell assume that the risk of 'Misappropriation without monitoring results' is 10% for SCI Foundation and the risk of 'False monitoring results' is 5%. Taken together this gives a total risk of something going wrong with your donation of 10%+5% =15%. GiveWell says that each dollar you donate to SCI Foundation is therefore only 'worth' $0.85 (ie 85% of $1). However, we can see this isn't a good assumption; if only 90% of your money ever makes it to the intended recipient (because 10% is misappropriated) then a 5% risk of false monitoring results means 5% * 90% of your original donation will be wasted, not 5% * 100% as GiveWell assume. This is only a difference of $0.005 per dollar donated vs the GiveWell assumption, but it matters a lot more when very large or very small percentage values are involved."

We see the logic of Froolow's argument, but we haven't prioritized addressing this because we expect it would make very little difference to our cost-effectiveness estimates.

We expect that the issue Froolow mentions here could change cost-effectiveness estimates by around 5%, but we don't expect that amount of change to be relevant to decisions about whether or not to make a particular grant in a large majority of cases.

Hard-coding in New Incentives and AMF CEAs

"There's a number of hard-coded formula in the New Incentives sheet (ie formula which contain an actual number rather than a reference to a cell), which therefore won't update when you make changes to the model. This occurs on row 80, 81, 146, 162, 170. There's also a stray hard-coded formula in AMF which occurs in Cell B28, which also contains an Inconsistent Formula Error, making it potentially the most erroneous cell in the entire model."

We have checked the hard-coded cells Froolow mentioned. We agree that it's not a best-practice to use hard-coding, but we don't believe that these cells introduce any errors into our CEA.

We reviewed this issue after reading Froolow's claim, and we don't think that any of these cells would cause any issues if other parts of the model were changed. The AMF cell, for example, is

just a calculation for LLIN lifespan in DRC. We'd want that number to stay the same, even if other inputs in the CEA were changed.