

GiveWell Donor Event
November 29, 2023

This transcript was automatically generated using software, with additional editing afterward, and may contain minor inaccuracies. If you have questions about any part of this transcript, please review the original video recording that was posted along with these notes.

Elie Hassenfeld: Hello everyone! I am Elie Hassenfeld. I'm GiveWell's co-founder and CEO and I'm so excited to welcome you to our year-end virtual event. Thank you so much for joining us. As many of you probably know, GiveWell is an organization that does research and recommends donations to donors around the world based on our research aiming to save or improve lives in low-income countries. We are really excited to be joined today by Emily Oster and others. As you probably know, Emily is the author of a book called *Expecting Better*. This is a book that I read right when it first came out, which was after the birth of my eldest daughter. But I was glad to know after the fact what all the tests that we went through during my wife's pregnancy were. We've had some other kids since, so it's also been helpful in those subsequent pregnancies. You know, I'm a big fan of Emily's. I'm also jealous of her because her running times just continue to elude me. I hope that someday I'm able to run a 10K with sub-seven minute miles. I don't know if I'll be able to, but Emily just sort of leads that target for me to be aiming at. In addition to Emily, we're joined by Svetha Janumpalli, who's the CEO of New Incentives, and Erin Crossett, who's a program officer at GiveWell. And today, we're going to generally talk about programs that focus on improving maternal and neonatal health and outcomes. Why are we talking about this today? Why is this a focus of GiveWell's right now and also this conversation? Well, the reason is that neonatal deaths, meaning deaths in the first month of life, account for a very high proportion of under-five mortality in low-income countries.

So collectively, about 2.7% of all newborns who are born die in the first 28 days. And these deaths account for 40% of all the child deaths that happen in the first five years. So this month of life is an incredibly crucial time where effective intervention can have a huge impact on life and

well-being. GiveWell has come to focus on this somewhat late in our evolution; we're, as an organization, 15 years old. And the reason it's taken us so long to get to this place where we can focus as much as we are now on these programs, is that they're not quite as straightforward to implement as some of the programs that we recommend otherwise. I think GiveWell is best known for our top charities, which often implement mass distributions of commodities. And these are not simple, but at least they're straightforward. So, for example, employing people to deliver preventative malaria medication door-to-door during the high season for malaria is something that we've supported via Malaria Consortium's seasonal malaria chemoprevention program. That's like fairly straightforward in the scheme of things. But maternal and neonatal programs require intense integration with an existing health system to be there to provide the services needed both before, during, and then after birth. And this is something that took us longer to get to.

One program that we've supported that Erin will talk a little bit about is a program called kangaroo mother care. In brief, it's a program that focuses on encouraging mothers of low birth weight infants to initiate skin-to-skin contact extensively and early, along with breastfeeding, and this program shows significant reductions in mortality. This program is very effective, but it took us a long time, multiple years, to finally arrive at an organization that we could support that was implementing this program. And so, because of some of its complexity, which Erin will describe, it's taken us longer to get here and find programs that we could confidently support.

So I want to just very briefly explain what we're going to be doing here. In a minute I'll turn the discussion over to Emily, Svetha, and Erin. Throughout, we would love it if you could ask questions via the chat in Zoom. Your questions will be visible to us, the panelists. They'll not be visible to all other participants. We love hard questions, so please feel free to ask anything that's on your mind. Whether it's criticism of what we're doing, something that you think we might not want to answer, we're happy to try, or even just basic explanations and clarifications about the topic and the material that we're going through. Feel free to ask anything about the presentation or anything GiveWell related. So

thank you again for being here. And now I want to turn things over to Emily to take the discussion forward and hear from Emily, Svetha, and Erin.

Emily Oster:

Thank you, Elie. I am so thrilled to be here. Erin and Svetha, are you leaving me alone? Turn cameras on. Hello. So I'm so thrilled to be here. And I'm so thrilled that you guys all joined us. So, my name is Emily Oster. I'm a professor of economics at Brown University, and I run Parent Data, which is a website that uses data to help people make better decisions about pregnancy and parenting. I write books about using data to make better decisions about pregnancy and parenting. And so I really, I really care about data—that's like my love language. And I think that's part of what has brought me to GiveWell, and part of why this is such a great synergy. You know, this is an organization that cares about evidence and cares about making decisions based on evidence. And that's hard to do. And it's hard to find. And so there is so much of kind of what I, the way that I think that decisions should be made that is reflected in this organization. So that is why I'm so excited to be here. I'm particularly excited to talk about maternal and neonatal health, because this is actually like part of where I started. So my academic work is, some of it's about developing countries. A lot of it is about health. I have work about neonatal survival in the US and trying to think about some of these issues in the context of the US, which unfortunately isn't as far from some of these lower-resource environments as we as we wish it were, I think. So I think many of these questions come up again and again, and thinking about how we can use evidence to make better decisions here and abroad is just so, so crucial. So with that, um. Svetha, Erin, it's so nice to see you. I would love to have you both just share a little bit about how you came to GiveWell, what you do, how it relates to maternal and newborn health. And, Svetha, I'm going to start with you.

Svetha Janumpalli:

Hi, Emily. It's so nice to meet you. I'm a big lover of your work and just your general approach to decision making. I first came across GiveWell about 12 years ago and was a consumer of GiveWell's recommendations. I have a background in impact evaluation, and was intrigued by this new wave of philanthropy that GiveWell was championing. I was very frustrated by how decisions were made by some donors and foundations,

because they often used superficial assessments or exaggerated stories of impact to make decisions. So this new approach resonated with me because it was transparent, rigorous, and based on evidence. And so in the early days of New Incentives, I reached out to GiveWell and I thought, well, there's no way we stood a chance of qualifying for GiveWell-directed funding. But today, about ten years later, New Incentives operates a program that provides small cash incentives to caregivers to increase childhood vaccination rates and reduce childhood mortality. We recently reached our 2,000,000th enrollment since starting the program in 2017, and since then we've encouraged over 29 million vaccinations.

Emily Oster: That's amazing. That is great. Can I just ask a follow-up, which is like, do you have a particular childhood vaccination that you think you've had the biggest impact, that people were the most resistant to, or you've had the biggest impact on?

Svetha Janumpalli: Probably when it comes to effect size, Penta vaccinations. But I think the biggest impact is on BCG because then we're encouraging people to initiate the vaccination schedule. And one big question we had was could we attract new vaccinators or could this only be a program to encourage completion. So we're really happy that it works in areas with low baseline rates, and that we can attract people to initiate vaccinations in the first place.

Emily Oster: Erin.

Erin Crossett: Hi everyone, and thanks Emily and Svetha. I've long been a fan of New Incentives even before my GiveWell days, so it's great to share the stage with you. And Emily, I have an eight-month-old, so you are a bit of a household name at this point. And speaking of names, my nephew's name is Finn, which is also the name of Emily's son. And so my brother and sister-in-law, they both read *Expecting Better*. And my brother and I like to joke that you entered his subconscious and influenced ultimately the naming decision. So thank you for that. I'm a program officer on the New Areas team at GiveWell. The new areas team is focused on identifying cost-effective programs, interventions that are new to

GiveWell, so not new to the world, and are outside of our other portfolios areas like malaria, water, vaccines, malnutrition. And maternal and newborn health is an important area for us, as Elie mentioned, due to high existing rates of maternal and neonatal mortality in low-income countries. And so just to really hit this point home, in 2021 alone, 650,000 infants died in the first 28 days of life in low-income countries. And that's around 1,800 deaths per day. And the chance of a woman dying during pregnancy is around 41 times higher in a low-income country than in a high-income country. But there exist these relatively low-cost interventions that can avert these deaths. And so we've made a couple of maternal and newborn health related grants, and we continue investigating other funding opportunities in this space. And so I'm really excited to be here today to talk a little bit about what we've been researching.

Emily Oster:

That's awesome. So I want to talk a little bit about evidence, because I think we all are people who like evidence. And sometimes evidence leads you to sort of places you didn't expect. And so I was thinking about what, in sort of my spaces, times in which evidence changed how I thought about it. And actually there are a few things in my parenting research, but there's one thing in my academic research. So, many years ago, I did a study in Nepal where we were trying to figure out whether providing girls with menstrual cups, menstrual products, but we were giving them cups, would change whether they went to school. And so we did this whole big intervention in which we gave girls randomized access to these menstrual cups, and we saw all kinds of interesting things about how they could learn to use them and so on. But what we found was that when we went into the data and we got these like school diaries, we had like girls fill out when they had their period, and then we got all the diaries of when they had gone to school, and we ran this whole intervention at the end, we realized they actually weren't missing very much school due to their period. So it turned out like this was not as big a, thankfully, in some ways was not as big a problem as we thought it was. The intervention had had some small effect, but kind of space of problems to fix was small, and so it sort of has always come back to me as kind of make sure you know what the problem is you're fixing before you try to go, try to go fix it with your randomized trial. But I'm curious,

maybe we can start with, with Erin, you know, is there a place where you got some new evidence and it really changed how you thought about something or how you thought about investing in something or supporting something?

Erin Crossett:

I think there's an interesting example. We got some new-ish research surrounding multiple micronutrient supplementation or MMS. And so I can sort of talk about how that updated us and how we're thinking about how our grantmaking and sort of research investigations are responding. And so the problem here is that, as many people on this call know, pregnancy requires a lot of micronutrients, both to ensure the baby and mother have sufficient nourishment, the baby's growing, and people with poor diets might be lacking in some micronutrients and potentially have a micronutrient deficiency. And so, in response, typical guidelines in many low- and middle-income countries is to recommend iron and folic acid supplementation for pregnant women. These are two micronutrients that are particularly important in pregnancy and for fetal development. But they're not the only important micronutrients. And so people who have been pregnant or have had children on this call might recall that, you know, taking prenatal, the dreaded prenatal vitamins with multiple micronutrients during their own pregnancies. And these are pills that include a whole range of vitamins and minerals. So things like vitamin A, vitamin C, zinc. So the question is, is it better to have all of these micronutrients, or is there some added benefit of having them relative to sort of the standard of care: iron and folic acid.

And so there was evidence that came out a few years ago that looked across multiple high-quality randomized trials in a variety of low-income countries and looked at birth outcomes for women who were randomly assigned to be in the multiple micronutrient supplementation group versus birth outcomes for women who were randomly assigned to be in this standard of care iron folic acid group. And they found that encouraging pregnant women to take multiple micronutrient supplementation instead of only iron and folic acid supplements led to better health outcomes and, in particular, reductions in low birth weight and stillbirth, and also post-neonatal mortality. And so this was an update for us. I mean, we know that multiple micronutrient

supplementation is quite cheap. And in particular, switching from IFA to MMS we think is also likely to be cheap. And therefore the intervention is looking really cost-effective in a whole range of geographies. And so we're really excited about it and are currently looking into potential funding gaps that GiveWell could fill.

Emily Oster: So let me ask you a question about that because, like there are many things in the multiple micronutrient supplement package. And one approach to this, if you sort of think about evidence, is to just say, well, they're all really cheap. And so like who cares which one it is? And another approach is to say, well, it would be even better if we could figure out like, what is it? Do you have a sense of what, what is like a plausible candidate?

Erin Crossett: Yeah, I think the short answer is no. So in the evidence, the common formulation was developed by the WHO and the UN called UNIMAP. And it's a formulation of I believe 13 different micronutrients. And so that version of MMS was the one that I believe is the most commonly used in this meta-analysis, the study I was referring to that came out a couple of years, but there is a little bit of variation in both the micronutrients that are included and also the dosage. And another important thing that I don't want to get too into the weeds here, but I was thinking about the exact dosage of iron. So the iron folic acid supplementation typically is 60mg, or 30 or 60mg, and it's 60 in areas with higher anemia burden, whereas MMS formulation is 30. And so one question was, well if you move from 60mg iron to 30, are you likely to see, would that cause adverse health outcomes? And the interesting thing that we're seeing in the evidence is no, that that's not happening. But I don't think that we yet have a good handle on what particular, whether there's a particular micronutrient that might be sort of carrying most of the health benefits or whether it's something about the interaction, whether including something like vitamin C increases the absorption.

Emily Oster: It's like sort of maybe you need the vitamin C for the iron take-up. Yeah, that's super interesting. Okay. I could talk for three hours about micronutrients, but Svetha, do you want to tell us a little bit, Is there a

kind of time in which something, you learned something new? The best part of research: tell me about learning something new.

Svetha Janumpalli: Oh, gosh. So many times we've had a lot of pivots in our history, but maybe I could walk you through a couple major ones. You know, we're very, our belief is we need to follow the data and make data-based decisions. So we start, you know, in 2014, New Incentives became the first recipient of what used to be called GiveWell incubation grants. So we explored applying the conditional cash transfer model to prevention of mother to child transmission of HIV services. So unfortunately, while this program was effective, we quickly learned that it wasn't scalable. We collected data from hundreds of clinics and learned that the rate of HIV positive pregnant women was lower than expected. So we decided on that basis to fold the program because, you know, the publicly available data simply wasn't matching what we were observing on the ground. So we then pivoted into looking at, okay, how can we apply this to a wider target population? So we started incentivizing women with at-risk pregnancies to deliver in health facilities instead of at home to reduce maternal and neonatal mortality. So this time, we, you know, a randomized controlled trial showed that these incentives were successfully increasing the rate of facility-based births.

So great. But there were three big problems. And these problems reduced the cost-effectiveness of the program much lower than we expected. And so those three problems were, first, the evidence base behind the effectiveness of facility delivery in reducing neonatal mortality had some major gaps. Second, even if we were to take that research, let's say at face value, some of the components that were in these services that were effective weren't consistently applied in the clinics in our operating context. And then, third, the rate of HIV-positive pregnant women in the program was expected to reduce at scale and this was a primary driver of benefits. So based on this evidence and the updated cost-effectiveness, we decided to phase out that program. But it's with those lessons that we then did a third major pivot into what we do today, which is to provide small cash incentives to increase childhood vaccination rates. And an independent, randomized controlled trial

showed that in areas where we operate, the program doubles the rate of fully immunized children against control areas.

Emily Oster:

It's such an interesting story because I think that when we, you know, I live like inside the academy, right? And we, people run randomized trials and we sort of think of and I talk about randomized trials as the gold standard and you can be certain of causality and so on. But there's this piece of taking our randomized trials to the world. And sometimes we refer to that as external validity. And say, you know, is the estimate that I get from this population externally valid for something outside? And then there's another thing which is like something can be really good in a trial, and then it can run into problems when you try to put it in the world. People, I hesitate to get into Covid, but this came up some in masking. Like the difference between the physical barrier and could it be used perfectly and what happens in the world. And I think in many of these spaces you're working in, there is that as well. Like you've got something that we know works for the science, and we've shown a randomized trial can work, but when you try to take it into the field, it doesn't work in a cost-effective way or it doesn't work with that population. And I don't know, Erin, I'm curious, like how in the things that you see, how much does that kind of last-mile implementation come up?

Erin Crossett:

Yeah, it comes up really often. So we at GiveWell, we're often in the position of looking at studies that find big reported effects on health, for example, and then having to think critically about, like you said, Emily, whether and to what extent those results would actually hold up in the real world, so outside of a research setting, in a different geography, or with a different target population. And so one interesting example that I want to talk about quickly is this program that we're supporting called kangaroo mother care, which Elie touched on briefly. And so the problem here is that babies who are born premature or low birth weight are very vulnerable, right? They're susceptible to infection. They have trouble regulating their body temperatures. And because of this, they die at higher rates than babies born at term or at an appropriate weight for gestational age. So kangaroo mother care, or KMC, is one way of addressing this. And it primarily is about prolonged mothers or caregivers providing prolonged skin-to-skin contact. But there's also an

emphasis on exclusive breastfeeding, breastfeeding support, and post-discharge follow-up. And KMC is seen as really a low-cost alternative to conventional neonatal intensive care. And so we have really strong evidence from multiple high-quality randomized trials that KMC reduces neonatal mortality by about 30% in low birth weight infants compared to standard neonatal intensive care. And so that sounds great, right? There's this really big problem. And the solution seems somewhat straightforward. And it's low tech. And there's great evidence that it works. But implementing kangaroo mother care in real world settings is quite challenging because unlike a study context, you know, there's not there's unlikely to be outside researcher involvement, maybe less oversight, fewer resources are dedicated to really getting the program right.

And so just to paint a picture here really quickly, you know, NICUs, newborn intensive care units, in low-income countries often suffer from staff shortages. So needing to train mothers on how to do KMC would stretch workers even thinner. And then you also need enough space in the hospital so that mothers can sit, lay down, have privacy for skin-to-skin. And so there's just a lot of behavior change that's required on behalf of caregivers and health workers. And so for these reasons, we knew it wasn't going to be a matter of just taking the evidence and these effect sizes at face value and assuming, okay, well, we'll see that the same effects persist if we fund this at a much wider scale. And so we make adjustments in our internal analysis to reflect our belief that, again, we'd likely see smaller effects than these studies show. And then the one really quickly, the last thing I'll mention is that, as Ellie mentioned, we finally found an implementer who could overcome a lot of these operational challenges, the research institute for compassionate economics in Uttar Pradesh, India. And we're also funding an evaluation of r.i.c.e.'s model to see whether the effects on neonatal mortality in this particular implementation setting are, if they exist, and if and we also think that this could inform future scale up of KMC.

Emily Oster:

I love it. Yeah, I mean I think it's in the space that we sort of worry about. Like, what do we do if we don't have data, and what do we do if we have data, but we're not sure that it's that it's kind of applicable to the world. I

mean, this is in some ways a question for you, Svetha. Is that like, if you don't have data, if there's like, how did you know to try your third thing? You did some things that didn't work. Like what, you didn't, was there data? Was it like, we're going to create our own data? Like what puts you on that path?

Svetha Janumpalli: So there was some data. And so when we don't know the full picture, we always try to say, well, let's go and find the best data possible and then start to take actions and then collect more data to try to triangulate information and build our understanding over time. And so even, this happens all the time in places where we work. So for example, in areas where we operate, we don't have reliable data on the number of infants born per year or the number that get vaccinated. And so both of those are really important. And so even though we don't have the ideal data, we want to consider, well, what's the best data we can get. And so to overcome this we conduct household surveys which give us, you know, reasonably accurate estimates of the percentage of children vaccinated. And so with that, we can use that information to take decisions about where to expand and, importantly, how long to operate. Because without that data, we wouldn't even know, is the program working, are vaccination rates increasing, and whether we should be planning for phase out due to insufficient impact?

Emily Oster: Yeah, yeah. I mean, it's interesting you're basically saying you triangulate, you sort of figure out the individual pieces and then you try to kind of put them together. And so that is, I guess that's how we do data. Yes. Okay. There are an incredible number of really good questions. And so, and we're going to have one last panel question, which is what is one thing you're excited about. And I will tell you, I will go first. Um, so I wrote *Expecting Better* now a decade ago and GiveWell is, 2007 is the like origin-ish and relative to sort of those times, relative to 10 or 15 years ago, people's interest in using data to make decisions has gotten so much greater. And I think that that to me is like a very exciting possibility, because I do really think that for many of the things we do in our personal lives, and especially for the ways that we spend our money, in philanthropy, spending it in places where it will matter is really important. And using evidence to figure out what matters, not using anecdotes, not

using how we feel, like that is the key. That is how we're going to get the most lives improved per dollar. And so the fact that people seem to care about data more than they used to, makes me excited and happy. Svetha, I'm going to turn to you. What makes, what are you excited about?

Svetha Janumpalli: Um, well, besides your next book and maybe a book about helping decisions for people trying to get pregnant. So I'm really excited about just, you know, achieving more scale with our program. You know, now that we've laid the infrastructure and we have an active presence at almost 6,000 clinics, I'm really eager to see what other causes of mortality we can address for the same targeted population at the point of infant clinic visits, so that could be reducing malnutrition or deaths due to diarrhea, without incurring additional overhead. And this may also enable us to operate in areas that were maybe previously not cost-effective when looking at, you know, working on infant vaccinations alone. So overall, you know, we're really excited to just build a giant cost-effective delivery machine to reach millions of children a year based on what the data says is most pressing and needed, and having an agile organization that can respond to those needs so that we're working on what is the biggest problem to solve.

Emily Oster: Love it. Erin, what do you got?

Erin Crossett: All right. So I talked about kangaroo mother care earlier and particularly kangaroo mother care when implemented in health facilities, which is called facility-based KMC. And I'm really excited about the potential to implement KMC outside of healthcare facilities and in particular, something called community-initiated KMC, which is basically when community health workers visit mothers with babies at home within 72 hours of birth, and provide mothers of low birthweight babies with information on again providing prolonged skin-to-skin contact, exclusive breastfeeding, etcetera. Again, we have really strong evidence from a randomized trial that community-initiated KMC finds similar reductions in neonatal mortality, and it also might have the added benefit of sort of avoiding a lot of these implementation challenges that I mentioned

earlier. So we're currently reviewing new evidence about its effectiveness and thinking about potential funding opportunities, which is exciting.

Emily Oster: I love it. All right, I'm going to bring Elie back. And we're going to do questions. Right, Elie?

Elie Hassenfeld: We sure are. That sounds great. We are. I hope I'm back and I'm excited to dive in. We got so many questions from folks. We are going to do our best to roll through them and then just keep asking away, because we'll have a record of them and we'll try to follow up with people after if we can't get your questions. So just ask whatever is on your mind. My first question, Erin, this is going to be to you, which is there are so many different local contexts and conditions in the places that we are looking to fund programs. How do we think about interventions like malaria, micronutrient supplementation, and everything when we're trying to support programs in different contexts. And I should say this is a question that came from Nick.

Erin Crossett: Yeah it's a good question. So a couple of things here. I think, this is really a question of external validity, or to what extent do we think that findings from academic evidence in one setting would translate to another geography. For multiple micronutrient supplementation and a number of other interventions we look at, we can find some comfort in knowing that the evidence is pretty robust. And what I mean by that is that the effect sizes or the large reductions we find in, you know, stillbirths, for example, or reductions in low birthweight are pretty consistent across geographies. But there are gaps. And so, you know, we have open questions about, again, the question of whether MMS is more or less effective than IFA in areas where there's a higher anemia burden. We can get at that to some extent with existing evidence, but there is sort of a demand for more, more trial evidence in high-burden areas. And then there's also, I think, a question of implementation. And so there's really implementation of a given program is going to vary considerably across contexts. And so I think that this is just something that throughout our grant investigations we try to really have a solid understanding of sort of like what, in the case of MMS, for example, like what the standard of care IFA program looks like, you know, to what extent do governments need technical assistance

as they shift from IFA to MMS? So something that we've heard a lot is that iron folic acid is often produced or manufactured locally. So it's a little bit, it's more challenging and more expensive to actually import MMS. And so that is just an additional kind of implementation and operational barrier that would need to be overcome in that country, but might not be relevant in another.

Elie Hassenfeld: Emily, how do you think about the same question? Right, we have these RCTs. They're often about particular populations, particular locations, particular times. And then applying that as general principles for people to follow wherever they are in their circumstances. How do you think about that challenge?

Emily Oster: Yeah, I mean, I think that there's, so this is actually a question I've worked on like the technical side of. So I have a, we did some work about sort of how you could think about taking an experiment and like basically translating it to some other context and arguing that, like, there's sort of an underlying data principle you could use, which is basically reweight the demographics, which isn't perfect. But if you sort of said, like this place is, if you want to port something from one location to another, you have to have a view that like the places are in some way similar, and if they're not similar at all, then you can sort of forget about porting it over. And if they are similar, then you may be able to sort of use some of the characteristics to try to sort of figure out, like, is this, to what extent is this population like my other population? But, you know, I think it is a hard question. And it also involves thinking about how likely is it that this intervention, that there's heterogeneity in the treatment effects of this intervention, which is like a very technical way of saying, sometimes the interventions seem like they should kind of work the same biologically for everybody, and sometimes they don't.

And sometimes you might think the implementation problems are the same and sometimes they aren't. And so like disciplining a little bit, what do you mean when you say I'm going to do it somewhere else. Like why would it be different? You know, the example I always give is like, you know, if your intervention is like hitting people in the leg with an iron bar and your outcome is do they get a bruise, like that's pretty portable

across spaces, right? Because that's just like how human anatomy works. And so thinking about how close is my intervention to hitting someone with an iron bar in the sense of like, how much is it likely to vary across people is kind of one piece. And then, you know, how much is the way that I implement going to be similar is another piece. And even looking inside the existing intervention to see, do you have any sense of how much the impacts of the intervention varied across groups? Because that gives you some idea about how much heterogeneity in treatment is there in general for this intervention.

Elie Hassenfeld: Right. Yeah. So I mean, I think a lot of what GiveWell does in many ways, all of us here are doing, is looking at rigorous evidence, trying to then apply it, utilize it, learn from it in what we do in the real world. And so a few questions came in about this process of running and then utilizing randomized controlled trials in formulating decisions about what to do. So, Svetha, this is a question for you. This is from Luke. You know, there's a lot of worry about randomized controlled trials in low-income countries that are essentially, quote, doing experiments on poor people. And, you know, you obviously, you remember this, New Incentives participated in a large randomized controlled trial that GiveWell helped support to try and determine the extent to which conditional cash transfers would increase immunization rates in Nigeria. I'm curious what you heard about this challenge as you were going through the period of the RCT being conducted. Is this, you know, what was the public perception of the trial that you were conducting or you were participating in? And how did people feel about it?

Svetha Janumpalli: Yeah. This is a really interesting question. Sometimes I think this is a little bit of a modern trap where we might feel this more externally. You know, we asked ourselves this a lot because we didn't want to be seen as, you know, just furthering experimentation without benefits. But oftentimes, you know, what we see is people really respond to evidence-based decision making. And so sometimes it's easy from the outside to think, well, okay, do we need all this evidence or are these vanity metrics? Are these truly necessary to implement? And what we found is people appreciate evidence-based decision making, because then you take out a lot of the politicking and interlocutors of how decisions are normally

made. And so we based that, and I guess, you know, a lot of it comes to your motivations. Why are you doing that study? Do you have an intention to scale if it works? But I'd also say that applies to monitoring, because when we look at, well, how do people feel about collecting so much data, validating that information, monitoring that services are delivered. We often hear from caregivers that they like it because otherwise these things were disappearing. You know, their benefits were getting stolen on the way. And so they appreciate having accountability in the process.

Elie Hassenfeld: Totally. And I think this is something, just as an aside, we feel too, at GiveWell. There's a lot of limitation to—and we got some questions about this, but hopefully we'll get to them—relying on the programs for which there is a significant body of evidence. At the same time, there's also, and I would say, a larger limitation to trying to make decisions without rigorous evidence or without high-quality data to rely on. That has a whole other set of problems that we know well and are glad to avoid. The next question I'm going to direct to you, Erin, about randomized trials and how this fits into GiveWell's search for new programs. This is a question from Robin. And the question is that, we're in the position to support programs for which rigorous studies have already been conducted. But how do we think about the trade-off between supporting new programs that have existing evidence already versus potentially funding evidence that could then, on promising ideas, that could lead us to fund additional programs in the future? You know, in this search for new ideas, how are we thinking about do we do both of those two, or we do do both of those? But how do you think about those different pathways of funding?

Erin Crossett: It's true that we are interested in, there are a lot of programs that we fund because like you said, Elie, we have like a very robust evidence base. We can be sure that, okay, this works. Let's go implement it. But then there are a lot of other programs where we're like, okay, there might be some evidence. It's not great. Our best guess is that this could plausibly be cost-effective, and this could be the type of thing that GiveWell wants to fund, but we're really uncertain about the actual effect sizes. And so we think that it's worth funding, for example, a randomized

trial, so that we can, like, really shrink that, we can resolve that uncertainty, actually, right. Like we would fund a nice randomized trial, and hopefully at the end of it we would have, okay, this is our best guess of the effect of a given program on neonatal mortality. And based on that, we can actually determine, assuming that was our key uncertainty, what the overall effect, cost-effectiveness of the program and ultimately make decisions about whether we should fund something and, if so, at what amount. And so I think a lot of the sort of the internal calculus here is driven by, you know, how cost-effective we think something could be under sort of a range of optimistic or pessimistic scenarios, and conditional on it looking good and being able to resolve some uncertainty through funding research. Then there's this question of how much additional room for more funding is there? Do we actually think that there's a real funding gap that GiveWell would be able to fill, and, if so, then that seems like it would be a great opportunity for our involvement.

Elie Hassenfeld:

Yeah, this is something, just to share with everyone here, that we've done a fair amount of over the last few years, and it's something that I think we'll do more of in the future. As an organization that relies on studies to make decisions, I think we're in a good position to think about what studies would most change our minds in the future. And so it is a major part of what we're doing today. So this is a question from Sarah. Emily, I'm going to direct it your way though. Erin and Svetha, I'm also interested in your thoughts on this. And you know, this also relates to applying study results to the real world. And I think here the question is, you know, often the conditions under which a trial takes place are just fundamentally different than the way things will work in the real world. I remember when we first looked at studies of malaria nets, there were some in which, the surveyors would go and literally like visually check if people were sleeping under nets at 5 a.m. and that, I think, gave some data, but also served as a reminder that you should put up your net every day. And that's not what happens in the real world. And so, you know, one of the questions we got had to do specifically with, you know, there are small-scale programs where health workers behave a certain way. Maybe they behave differently under real-world conditions. But maybe more generally, Emily, like, how do you think about this problem, that

the conditions under which the study is conducted may be very different than the conditions that exist in the real world when a program is implemented at large scale?

Emily Oster:

Yeah, I mean, this problem has many pieces. So one is the sort of standard Hawthorne effect problem, which is what you're sort of what you're referring to. Like kind of the checking at 5 a.m. is an extreme form of the Hawthorne effect, where basically like the idea is like just by watching someone, their behavior changes because they are being watched, and then they're thinking about it, and that's going to drive your effects, could drive effects even if nothing is happening. And then there's the second piece, which we will often run experiments in places we think, not that we think that they're going to work there, but just by virtue of where it is easy to do interventions, they tend to be places that are functional. So you see this not, you see this in the US as well, so the the sort of largest trial of inductions to ask the question of whether labor inductions lead to more C-sections is a trial called the ARRIVE trial, which sort of came to the conclusion that labor inductions don't raise, you know, at or after 39 weeks, don't raise the risk of C-sections. And it's a really good trial. It's really well run. But one of the criticisms that's come up about this is, you know, the C-section rate in the hospitals in the trial is like 12% or something, or the C-section rate in the trial is 12%, the C-section rate in the US, like 35%.

And so obviously these places are different in some way. And it may well be the case that the way that the doctors were approaching induction or not induction is just different in these settings than it is elsewhere. And so you kind of often have that piece of it, just that the space of the intervention is different than the outside. And that's hard because ultimately it means the way that we have to evaluate evidence is sort of take a trial as suggestive. And then you can either, sort of similar to what I think Erin was talking about, kind of say, okay, well, I know I'm not getting there, let me shade down and it's still going to be worth it if I get 75% of the effect or half of the effect. So that's like one option. And another is to say we're going to run it small, we're going to run it medium, we're going to keep going up and big and sort of see how things are varying. Of course, once you're experimenting at scale, you're

doing it at scale. And so there's just this inherent tension between learning and doing, I think.

Elie Hassenfeld: Svetha, how have you thought about this question? You're operating an organization. You have trial results that show that it worked exceptionally well at increasing immunization rates. And now you're just operating. So how do you think about whether that impact that we observed in the trial is persisting?

Svetha Janumpalli: Yeah. Think this is what we obsess about 24/7. So when we look at the trial, it was based on, you know, self-reported vaccination behaviors, but also in the trial, you know, to minimize spillover, it selected one clinic, let's say, out of a cluster of many. So in the first case at scale, we need to make sure that reported vaccinations actually translate to vaccinations administered. So to make sure that that translated to the real world, we had to implement a lot of different procedures at scale. So looking at child health cards or immunization cards, looking at aggregate data on the number of vaccinations administered, and then making sure we're doing direct observation of vaccination, because otherwise we ran the risk of just having a program that incentivized the recording of vaccinations, but not the administration. And then when it came to the setup of clinics at scale, for fairness and reach we wanted to be able to operate in all clinics so that we didn't cause distortions in the health system. So that was a big business model costing question. And so we had to figure out, well, at scale, how do we do that cost effectively. And so it was kind of iterating and testing as we started to scale up the program that we figured out how to operate at all clinics, which ended up being a lot more beneficial in reaching all caregivers, but was difficult to work out in the beginning.

Elie Hassenfeld: Yeah. That's great. Thank you. And I guess at a high level, I think this challenge of interpreting rigorous evidence and then using it to make decisions is central to the GiveWell project. We recognize it's certainly not as easy as saying a study said X, therefore that is the answer everywhere and in every case and it's simple. In fact, it's very complicated. And a large part of what we see as our value added is trying to understand the way in which evidence can be applied to the

real world, and then following up after we've supported organizations like New Incentives to say, how is this actually working? We have a lot of interesting stories we can share about places where later data has told us that the initial results didn't hold, and we had to go a different direction, which we're happy to share if anyone's interested, another time. Emily, wanted to come back to you with a very different type of question. This is a question from Kara, and it's about stimulation as important for early childhood development, for baby's development. And the question is, do we know anything about the minimum dosage for impact on stimulation for babies?

Emily Oster:

Not really. I mean, you know, I think this is a complicated space because we know that, like, leaving a baby alone in a crib with no adult, you know, even if their basic needs are met, is definitely not okay and is, like, awful. We also, I think, know that the advice, like you must narrate your entire day to your child and tell them every single thing you're doing and explain every time you change their diaper, like what you're doing and where the day, and like that, that is also unnecessary and that babies can be left alone for, you know, some periods to sleep or entertain themselves. And so, you know, I think it's an example for me, actually, of where, I think, we end up with a pretty large disconnect between the kinds of pressures that the parents I often talk to feel they are under and the reality of what is important. And so, you know, the years from 0 to 3 are super, super important in kids' development. But that is different from saying that, like, you have to only use wooden toys.

And there's somehow I think we miss a little bit the idea that, like protecting kids from toxic stress and making sure they have a nurturing caregiver where they feel safe and enough to eat and health care, those things are really, really important. And most of the rest of it, is kind of a little bit here and there. And so I think that, you know, for me, maybe this isn't quite an answer to the question about stimulation, but I think it's an answer to how I would ask that question, which is like, what do we need to get to a place where people can be with their kids, and their kids can have a safe place and enough food and enough sleep, and like you know, somebody who's telling them they love them and hug them, that's really important. And then further stimulation, you know, do we need

flash cards that have different colors on them or only black and white, or do you need to, you know, get the right mobile and whatever other things people are thinking about mostly is just, don't think about that anymore.

Elie Hassenfeld: Yeah, great. Thanks. I'm glad. I'm glad that that's the case, because otherwise I've done a terrible job as a parent.

Emily Oster: No. Totally. I mean, it's an interesting space for me in working with, sort of simultaneously speaking to a lot of people for whom they really are trying to spend a lot of time thinking about what's the right, you know, makeup of particular toys. And then, you know, having done a lot of work in the space, you know, in developing countries and in maternal mortality, in those spaces, and say, you know what? Don't worry, it's fine. Don't worry about it.

Elie Hassenfeld: Yeah, great. That's great. I want to ask, and then I guess I'll answer this question that came from Megan. It was just asking whether GiveWell is planning to look into reproductive health programs, programs that focus on contraception specifically. And yeah, I mean, the answer here is, you know, first of all, well, we do, we would like to look into it. I think it's important to know about GiveWell that we know that we don't have all the answers. We haven't covered everything that's out there that is promising. You know, we've been around for about 15 years. We have a research staff of 40. We're covering more ground today than we ever have. But also, you know, there are practical limits to what we can look into in the short term, and we're always making challenging prioritization decisions about what to focus on. And so for various reasons, reproductive health has been a complicated area and one that we have put a little time into and supported a few programs historically. It's one where we really hope to be able to spend more time on it in the near future, and we hope that a lot of the expansion that we've done in the last few years, both in terms of staff size, we have more researchers who can do more, but also in being more open to a more complex set of programs, and being more capable at evaluating them, is something that will enable us to look into a very wide variety of programs of which reproductive reproductive health is one.

I wanted to ask, so another question that came in is from Henry, and Erin, I'm curious, you know, if you want to answer this or pitch it back to me if that would be helpful. But I'm curious how you think about, imagine program A, which has a lot of uncertainty in the evidence, but it could be really, really high impact, and then on the other hand, you have very, very certain evidence or as certain as it gets. But, you know, the confidence intervals around the possible impact are much narrower. So higher certainty, lower impact, lower certainty or more uncertainty and higher potential impact, I don't know, how do you take that into account when you're thinking about new programs to investigate?

Erin Crossett:

Yeah. So I think a couple of things here, I think we'd want to know whether, again, whether and to what extent we could resolve some of the uncertainty. So the program, I can't remember A or B, the one with wide confidence intervals, like, is there a way that GiveWell's funding could resolve some of that uncertainty, especially if the upside potential is high. Right. We're talking about a program where we're like, we're not totally sure if it works, but if it does, it could be a real game changer. And then ultimately it comes down to cost-effectiveness, right? So I think we're focusing, we're talking about here the benefit side of the equation. But it's also important to think about the cost. So for the program that's like a little bit, you know that could have potentially much higher returns but is very uncertain. You know, is it extremely expensive to implement? In which case the cost-effectiveness might pencil lower than the one that may be a surer bet. But ultimately, you know, it just comes down to the details there. But I'm curious if you would add anything to that.

Elie Hassenfeld:

That's a great answer. We're coming down to our last question. So yeah, no that's great. I agree with that. The last question, is it sort of asked from, related to GiveWell. But I think it's a good question for everyone here. So maybe, Emily, interested in your take first. And then we can sort of hopefully hear from everyone briefly. Which is, you know, strong focus on evidence-based programs or evidence in decision making is obviously very limiting. You know, the world of things that have rigorous evidence or the universe of places that have rigorous evidence is very narrow compared to all the possible choices that are out there for GiveWell, but

also in real life. And I'm just curious how, you know, each of you, maybe briefly, thinks about that in the work that you do, the limitations that come from evidence, and I don't know, how have you decided to focus so heavily on rigorous evidence in what you do?

Emily Oster: I think almost always evidence is only a piece, but I think that it provides us a framework from which we could evaluate what other things we want to think about. So often you will say, you know, the data doesn't boss you. The data tells you, you know, this thing is more effective than this thing, or this has this small effect in this direction or not. But there could be something else. You could say, I don't want to fund that program because I'm ethically opposed to it, or I don't want to do this parenting thing because it doesn't work for my values. And so data is not the answer, but it is something we should always be starting with, or at least thinking about first before we move to adding these other pieces. Yeah.

Elie Hassenfeld: Svetha or Erin, either of you want to jump in?

Svetha Janumpalli: What really draws me to evidence and evidence-based decision making is accountability, because it seems like otherwise we can be operating or doing something for a long time without having the feedback to course correct or know if something's working. So that's what really attracts me is, you know, the whole system ticks and is kept accountable with that lens.

Erin Crossett: The last thing, I think I agree with everything that Emily and Svetha have said, the last thing I'll add is just, you know, these issues are really important. We're talking about, you know, maternal and neonatal deaths here. And so from my perspective, I think the evidence just makes this problem much more tractable than talking about things in sort of like a broader, more macro sense. And it helps us, from my perspective, identify, again, particular strategies to, in this case, you know, avert infant and maternal deaths, which I think is the ultimate goal here.

Elie Hassenfeld: Yeah. That's great. Thanks, Erin. I was thinking I might add something, but I just agree completely with what you all said, and so just want to echo it all. We're going to close. I wanted to thank you, Emily, Svetha,

and Erin, for joining us and participating in this panel today. It was great. It was incredibly interesting and informative. And also want to thank all of you for joining from around the world to participate in this today during what I imagine is a busy work day for nearly everyone. It's amazing to me. GiveWell started more than 15 years ago. At the time, what we were told is no one would care at all about evidence and impact and philanthropy. That's not the way people give. And, you know, more than 15 years later, you are helping us raise and direct hundreds of millions of dollars a year. Hundreds of you are joining for a wonky discussion of evidence to help people in low-income countries. It's amazing the change that we're seeing is outstanding. And we're so, we're so grateful for your support and your participation in our work. So thank you very much and have a wonderful rest of your day.