Journal of Benefit-Cost Analysis (2024), 15: S1, 127–154
doi:10.1017/bca.2024.10

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Lessons from Applying Value of Statistical Life and Alternate Methods to Benefit–Cost Analysis to Inform Development Spending

Alice Redfern[1], Sindy Li[2], Martin Gould[2], Felipe Acero[3] and Daniel Stein[4]

[1]IDinsight Philippines, Manila, Philippines
[2]Open Philanthropy, Berkeley, Berkeley, CA, USA
[3]Independent
[4]IDinsight Inc., San Francisco, CA, USA
Corresponding author: Alice Redfern; Email: alice.redfern@idinsight.org

Keywords: choice experiment; development; Ghana; Kenya; LMICs; stated preferences; VSL; WTP

Abstract
Estimating value of statistical life (VSL) is an important input to many benefit-cost analysis (BCA) approaches, but for many low- and middle-income countries, there are limited or no data estimating VSL. Current guidance relies on extrapolation of results from high-income settings, which may be unreliable, leading to low confidence applying VSL. During 2019, we surveyed 1,820 low-income individuals (average consumption per capita USD329) across four diverse regions in Ghana and Kenya, to inform recommendations about effective spending in the development sector. We elicited VSL using a stated-preference approach, capturing the willingness-to-pay to reduce the risk of death for themselves and their children. Additionally, we conducted multiple "policy choice experiments" (PCEs) in which we asked respondents to choose, from the perspective of a decision-maker, between programs that save lives of different ages, and save lives and provide cash transfers. VSL estimates for this population fell in the range of USD66,795–USD90,453 (PPP-adjusted). We found similar results in the PCE but uncovered much stronger preferences for saving younger lives. Overall, our results suggest that VSL in low-income countries may be higher than estimates based on extrapolations from wealthy countries and that within these communities, policymakers should place more weight on saving the lives of young children. We also explore methodological learnings about how to apply and collect data for BCA in particularly low-income, low-education settings. We find that through careful training and gatekeeping, it is feasible to elicit complicated preferences in this population, and both approaches have their benefits and drawbacks.

1. Introduction

Many important decisions involve the trade-off between monetary and health outcomes. At the personal level, someone may choose whether to pay for a better and more expensive medical treatment, take a riskier and higher-paying job or a riskier and faster means of transportation. Similarly, at a societal level, governments face these trade-offs when setting pollution standards, speed limits and allocation of resources to health versus other programs.

To navigate these choices, governments and other decision-makers often turn to the concept of the "value of statistical life" (VSL). VSL enables policymakers to quantitatively weigh the monetary worth individuals place on small reductions in their risk of premature death. By combining VSL with impact estimates of the relevant programs or policies, resource allocation decisions can be informed by how relevant populations make these challenging trade-offs.

However, employing VSL as a tool to inform the balance between monetary and health outcomes comes with certain complexities. One significant consideration is the availability of high-quality data to inform VSL estimates of populations in low- and middle-income countries (LMICs). Without VSL studies of appropriate populations, resource allocation decisions will fail to reflect the preferences of the individuals impacted by the relevant programs and policies. Furthermore, standard VSL elicitation approaches are based only on individual choices and therefore do not incorporate additional factors beyond the individual, such as equity or fairness, that might also be important to individuals affected by a given policy. Therefore, relying solely on VSL could potentially result in decisions that do not fully capture the richness of public perspectives.

Our study aims to address both of these challenges. First, we conduct a VSL survey with low-income populations in Kenya and Ghana, using a stated choice method popular in high-income countries (HICs). We gather VSL for both adults and for children of different ages. With the same participants, we also implement a novel "policy choice experiment" (PCE) in which they allocate public resources between a cash transfer program and life-saving interventions. This PCE has a number of attractive features compared to traditional VSL elicitation: it eliminates issues with individual-level liquidity constraints, understanding of small probabilities, and has the ability to incorporate respondent preferences beyond the individual, such as distributional concerns. On the other hand, it deals with a specific policy choice, and therefore unlike VSL the responses are not applicable to a wide array of policy considerations. The PCE may introduce other sources of bias; for instance, it relies on a respondent "switching" their choice between interventions as their ratio changes, but the estimation model performs poorly when participants never switch. By comparing challenges of field-based elicitation as well as implied VSL between the two methods, we test whether a PCE provides significant improvements over a traditional VSL approach when considering a specific policy choice. We provide additional color to the results by collecting qualitative data on why respondents made certain choices.

1.1. Existing VSL estimates from LMICs

While VSL estimation studies are regularly conducted in HICs, and study results are used in HIC policymaking, such studies are much less common in LMICs. This is likely due to a variety of reasons, including fewer resources for such studies, challenges collecting the survey data for stated preference studies and a lack of observational data (such as employment, salary or consumer data) for revealed preference studies.

This has led to a significant gap in the VSL literature. In review of the existing literature, Robinson *et al.* (2019) identified only 26 studies in LMIC contexts. Of those only 5 were conducted in lower-middle and none were conducted in lower-income countries (the remaining 21 were in upper-middle-income countries), and only 2 of those studies attempted to capture

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

VSL in Africa.[1] The review also highlighted the variable quality of published studies. Since then, a small number of additional studies using different approaches to measure VSL have been published in low-income countries (most notable for their similar target populations to this study are Patenaude *et al.,* 2019, and Trautmann *et al.,* 2021). However, most Sub-Saharan African (SSA) countries still have no primary data on which to estimate VSL.

1.2. Benefit transfer approach to estimating VSL for LMICs

In the absence of empirical studies and given the variable quality of studies in other LMICs, researchers and policymakers have relied on approaches that extrapolate values from HIC estimates, often based on the per capita gross domestic product (GDP). For instance, Robinson *et al.* (2019) provided guidelines for transferring VSL to LMICs based on the best available evidence. This approach has facilitated the publication of several benefit–cost studies of policies and interventions in LMICs, employing a consistent methodological approach (e.g., Radin *et al.,* 2020; Watts *et al.,* 2022; Syuhada *et al.,* 2023). Prior to this, similar guidance was given by Viscusi and Masterman (2017), which took a similar conceptual approach to the transfer but used differing elasticity assumptions and based their base VSL on a different sample of U.S.-centered studies. Alternatively, many researchers lean on WHO guidance, which suggests that an intervention is considered cost-effective if it saves one statistical life year for a cost of between 1 and 3 times per capita GDP of the relevant country (despite more recent attempts by WHO to distance themselves from this threshold, see Kazibwe *et al.,* 2022).

These extrapolations rely on an assumed relationship between VSL and income. While there are some data on the income elasticity of VSL (Hammitt and Robinson, 2011), due to the lack of data in populations with significantly lower incomes, it is not clear if these assumptions hold when applied to LMICs. This leads to high uncertainty around VSL estimates for LMIC populations, meaning that results are unlikely to be sufficient for policymaking and informing other large-scale resource decisions. This uncertainty is highest for populations living in extreme poverty who are most commonly missing in existing VSL data, and yet also commonly impacted by resource allocation decisions in the development sector. As a result, Robinson *et al.* (2019) note that "more research on the value of mortality risk reductions in LMICs is essential." Our study directly contributes to addressing this gap in the literature and represents, to the best of our knowledge, the first stated preference study of VSL in both Kenya and Ghana.

1.3. Methodological challenges to eliciting VSL in LMICs

In any VSL study, researchers must be aware of the inherent limitations of stated preferences approaches. This includes hypothetical bias (where individuals' willingness-to-pay (WTP) values do not match what they would do when faced with the same decision in the real world), misunderstanding of small probabilities leading to scope insensitivity, and assumptions around rational decision-making and proportionally to risk. In a population with very low household income, and limited access to financial services (i.e., a strong liquidity constraint), special efforts need to be taken to capture true underlying WTP values.

[1] A revealed preferences study in Tunisia (Benkhalifa *et al.,* 2012) and a single stated preference study in an urban setting in Sudan (Mofadal *et al.,* 2015).

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

As detailed in Section 2 below, our study aimed to reduce the impact of these limitations in our context through various approaches. We implemented a training and testing module based on visual aids and scenario-based questions, to familiarize respondents with probability and mortality risks and allow us to quantify comprehension levels (which we then used to in our analysis). To address the impact of liquidity constraints on respondent answers in a low-income setting, we asked for WTP over a 10-year period in manageable installments.

1.4. Relative value of mortality risk reduction for children versus adults

Many international development interventions working with people facing extreme poverty focus on improving the health and reducing the risk of death of young children. For example, malaria is a major killer of children under 5, and interventions such as long-lasting insecticide nets and seasonal malaria chemoprevention have been shown to reduce infant and child mortality (see Pryce *et al.,* 2018; Cairns *et al.,* 2021).

There is even less evidence available for LMICs to determine the relative value to place on mortality risk reduction for children. Many standard approaches to estimate VSL are even more challenging (revealed preferences studies relying on wage data do not apply, and young children do not have the competencies to answer complex stated preference surveys). Most available estimates for HICs rely on stated preferences of their parent, or averting-behavior data (e.g., WTP for bike helmets). A review of available VSL studies (Robinson *et al.,* 2019), entirely conducted in HICs, has found that the range of values for children captured range from 1.2 to 3 times those captured for adults. A midpoint value for children of approximately twice the value assigned to adults is recommended for use in policy decisions. A recent systematic review by Peasgood *et al.* (2024) on the relative value of health for children compared to adults similarly found children are valued higher (and that this is relatively consistent across different measurement approaches used). To the best of our knowledge, only one prior pilot study captured a ratio of child to adult VSL in an LMIC (Bangladesh),and found no difference (i.e., ratio of 1.0) (Odihi *et al.,* 2021).

Given the importance of valuing outcomes related to children, and the scarcity of data available on this, our VSL survey included questions for participants on their willingness to pay for small risk reductions affecting their children. We also implemented an analogous PCE that assessed the relative importance of averting risk of death for different age groups.

1.5. Alternative estimation approach: policy choice experiment

Besides traditional VSL elicitation, we also include a novel PCE, in which respondents are placed in the position of a policymaker deciding between allocations to different programs: an intervention that saves lives and cash transfers. Our PCE is conceptually as close as possible to the type of resource allocation decision faced by a decision-maker allocating resources in international development, and is in fact based specifically on the type of funding allocation decision faced by our partner in this research, GiveWell. This framing is similar to the “societal perspective” framing that has been used elsewhere to measure health and non-health benefits or prospective treatment (such as Kwon *et al.,* 2017; Sheen *et al.,* 2023). The framing has also been used elsewhere to assess “ethical preferences” toward the relative value of averting deaths of individuals of different ages (e.g., Johansson-Stenman *et al.,* 2011; Palanca-Tan, 2013).

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

There are a couple of reasons why the PCE is a useful complement to VSL elicitation. The first reason is the aforementioned methodological challenges with capturing VSL. Capturing stated preferences about VSL relies on good respondent comprehension and acceptance of the presented scenarios, including understanding of small probabilities. Additionally, an individual's willingness to pay in any hypothetical scenario will still be closely tied to their ability to pay. Both of these constraints are particularly apparent in a target population with extremely low levels of literacy and low liquidity. This may lead any decision-maker to question the validity of VSL results captured in this population. Our choice experiments place the respondent in the perspective of a decision-maker for their community, removing any personal liquidity constraint. They also deal with risk in real terms (i.e., numbers of individuals affected in a community rather than a small change in risk to oneself), which can make the concepts easier to grasp. While VSL may be the standard approach, triangulating the results with an alternate measurement even if less comparable to the literature, can serve to increase overall confidence in the results.

The second reason for including alternate approaches is to capture preferences that go beyond individuals' personal trade-offs. The direct choices that decision-makers in governments or donors face are generally different from that faced by individual families. Donors generally must decide how to allocate resources among a number of programs and across a wide population. Traditionally elicited VSL can be an important tool in these allocation decisions specifically allowing the comparison of interventions with monetary outcomes to those that impact mortality risk. However, when translating VSL into a full policy decision, there are other potential inputs that may be included in determining the final choice, such as fairness, customs, distributional concerns and so forth. When asking the participant to take the perspective of the decision-maker, they may already consider and incorporate these factors into their preference in a way that they would not when just considering their personal trade-offs.

However, the PCE approach also comes with potential downsides. First, it may introduce its own biases in elicitation. For example, the PCE elicitation approach asks participants to trade-off between a number of cash transfers and a number of lives saved. When speaking about actual lives saved may be easier to understand than decreasing risk (as in traditional VSL elicitation), it may also introduce bias as participants may feel inclined to prioritize saving lives with certainty. Relatedly, the PCE approach is estimated based on respondents "switching" their choices between lives saved and cash transfers, as the ratio of beneficiaries between these programs changes. When respondents switch their choice between lives saved and cash transfers, this allows one to bound their preferences, allowing one to derive an estimate comparable to VSL. If respondents always choose to save lives (no matter the amount of cash transfers suggested), then the elicitation does not provide an upper bound on the monetary value of life for that respondent. As the number of respondents who do not switch increases, the model's aggregate estimation becomes less reliable.

Finally, another drawback of PCE compared to VSL is that it deals with a very specific policy choice, while VSL is a general concept that can be applied to a wide array of policy decisions. Therefore, if one believes that the PCE is a more reliable estimation approach, a policymaker would want to do a separate elicitation for different policy decisions. This makes it less practically useful.

Our PCE is designed such that it can be used to extract an implied VSL. We therefore can compare the directly elicited VSL from the implied VSL from the PCE to understand if the approaches lead to different policy implications.

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

1.6. Additional study context

This study was a collaboration with GiveWell (GiveWell, n.d.), a non-profit "dedicated to finding outstanding giving opportunities" through "in-depth research to determine how much good a given program accomplishes (in terms of lives saved, lives improved, etc.) per dollar spent." Our research aimed to inform the quantitative value GiveWell places on different good outcomes (such as saving the life of a child, saving the life of an adult and increasing the consumption of a household). This required exploring and testing a variety of methods for measuring the VSL and similar concepts in populations experiencing extreme poverty. Ultimately, we sought to understand how the preferences of program participants could be integrated into GiveWell's resource allocation decisions effectively.

Finally, to provide rich contextual information with which to interpret our VSL and choice experiment estimates, we also used several secondary methods to capture data on the following:

(i) the rationale and moral reasonings used by beneficiaries when making trade-offs;
(ii) the subjective well-being of beneficiaries (as measured by self-reported life satisfaction), including how this correlates with different individual characteristics;
(iii) other information about beneficiary lives, including primary data on the indirect effects of death (economic and emotional), and secondary data analysis on the economic contribution to the household by age.

While space precludes this paper from covering these elements in-depth, the interested reader can find further information in an online report covering this project (IDinsight, 2019).

2. Methods

2.1. Sampling approach

We gathered information on VSL using a stated preference approach, using a survey conducted in Migori and Kilifi Counties, Kenya and Karaga and Jirapa Districts, Ghana from May to September 2019. The goal of our sampling approach was to identify a quasi-representative sample of households living in poverty with diverse demographics.

Within each country, regions were selected purposively to ensure that data collection was conducted in areas with high rates of poverty and mortality of children under the age of 5, and with geographic and religious diversity. Within each region, we used probabilistic sampling to select smaller geographical units: sub-county (Kenya) or electoral district (Ghana), and then rural village (80 % of sample) or urban community (20 % of sample). Within each village or community, we identified eligible households living in poverty by first conducting a Participatory Wealth Ranking exercise (Wiegand, 2020) with recognized community leaders and then a verification listing exercise using the Poverty Probability Index.[2]

In total, 1,846 respondents from identified relatively low-income ("poor") households across selected communities were interviewed, in addition to 246 relatively high-income ("wealthy") households to allow for validity checks and comparisons of our results across

[2] The combination of PWR and PPI has previously been used by the charity Village Enterprise to target poor households in Kenyan and Ugandan communities similar to those in our sample (see Doty, 2014).

income levels. These "wealthy" respondents were not included in the final estimates in line with our overall sampling approach and study objectives targeted at households living in poverty.[3] In order to allow comparisons with other studies based on household wealth levels, we estimated household consumption using a consumption module adapted from Haushofer and Shapiro (2016). To obtain gender balance, within each household, we randomly selected whether a male or female adult respondent should be surveyed. Ninety-four per cent of respondents we approached consented to and completed the survey.

2.2. Contingent valuation design

In designing our VSL survey questions, our primary goal was to produce an instrument that respondents in our context would understand and be able to relate to. The secondary goal was to maintain a degree of consistency with other stated preference VSL studies, especially those in LMICS, to allow for comparability of results. We drew heavily on the question design in Hoffmann *et al.*'s studies in Mongolia (2012) and China (2017) and were informed by an extensive review of literature on the effective communication of probabilities and risk (notably the reviews by Corso *et al.,* 2001; Garcia-Retamero *et al.,* 2012). We adapted our survey questions and visual tools through extensive piloting first in Kenya throughout 2018, then in Kenya and Ghana between January and May 2019.

An in-depth summary of our survey contents can be found in Appendix A. Our survey started with a training and testing module, adapted from a version trailed in urban Bangladesh by Mahmud (2011). Respondents were introduced to the concepts of probability and mortality risks through visual aids containing a series of scenarios.[4] For each scenario, respondents were asked which chance or risk was higher or lower and if the respondent gave the wrong answer, a scripted explanation was given and the number of explanations required to reach the right answer was recorded.

After completing the training and testing module, individuals were introduced to the WTP scenario. Interviewers introduced respondents to the idea that there was a new disease affecting their community that kills a small number of people who catch it (20 in 1,000 over 10 years). We then explained that a health product[5] (randomized to be either medicine or vaccine in Kenya, medicine in Ghana) has been developed which reduces the risk of dying by 5 in 1,000 or 10 in 1,000 within the next 10 years. Respondents were asked for the maximal amount they were willing to pay for this hypothetical health product to reduce their own risk of dying. If the respondent was a parent or primary caregiver to at least one child in the household, a child in the household was selected at random and the respondent was also asked for the WTP for the same hypothetical health product to reduce the selected child's risk of death.

[3] Neither the sampling or analysis was designed to include the wealthier respondents, and these respondents did not complete all elements of the survey (only those necessary to validate results).

[4] The scenarios, detailed in full in Appendix A, were (i) different chances of winning a lottery, (ii) roads with differing risks of experiencing a car crash and (iii) individuals with different risks of dying over the next 10 years.

[5] In Kenya, the health product was randomized between medicine and vaccine to allow us to test for potential framing effects. In Ghana, we only used medicine as during piloting we found there was not a clear enough distinction between the words for medicine and vaccine in the two local languages in which the survey was conducted.

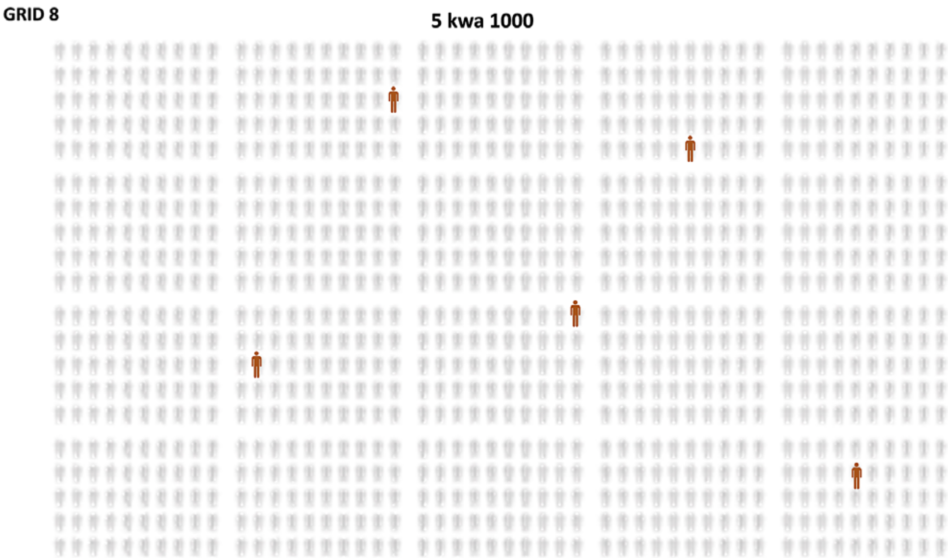

Figure 1. *Example of visual aid used to represent mortality risk. In cases where mortality reductions were visualized, green color coding was used to represent individuals with reduced risk and red for those who remain at risk.*

A visual aid was used to represent mortality risk levels (see Figure 1). The following elements of the elicitation were randomized: type of hypothetical health product (medicine or vaccine), the age of the child selected, the order of the risk levels presented and whether the individual was asked about risk reduction for themself or for their child first. To elicit WTP, we used a payment card (a visual tool presenting a list of price options that the respondent can choose from, adapted from Hoffman studies). However, during piloting, we found that the payment card was not effective for the relatively large proportion of the sample who were illiterate. In practice, therefore, the payment card was used as a visual tool to support enumerator explanations and respondent decision-making, and any value was accepted as the respondent's answer even if it was not represented on the card. Given the extreme liquidity constraints experienced by the individuals in our sample, we first asked for an initial WTP, then asked for WTP if they could pay for the health product over 10 years (the same time frame of the risk reduction) in small installments of their desired frequency (monthly or annually) (Patenaude *et al.,* 2019; Trautmann *et al.,* 2021), notably applied a similar approach allowing repayments in a similarly poor population.

Our question design allowed us to identify respondents with poor comprehension of small probabilities and our WTP scenarios. In Section 3, we present results where respondents with poor comprehension are dropped. In Appendix C, we include results from the full sample for comparison. The survey design also allowed us to test whether individuals were sensitive to the size of the risk presented (i.e., were individuals willing to pay more for the higher risk reduction, known as the internal scope test). We also randomized the order in which the two risk reductions were presented allowing us to assess if the sample population demonstrated overall sensitivity to the size of risk presented (i.e., is the population presented the 10 in 1,000 reduction first, willing to pay more than the population presented the 5 in 1,000 reduction first, known as the external scope test). Finally, we examined the relationship between WTP and respondent demographics, including household consumption, reported life satisfaction, health status, region, religion, age and gender.

To estimate VSL, we take WTP and divide by the risk reduction level. To allow for direct comparison to recent studies conducted in similar populations in Tanzania (Patenaude *et al.,* 2019) and Burkina Faso (Trautmann *et al.,* 2021), we also estimate value of statistical life year (VSLY) by dividing VSL by the expected life years remaining for the respondent using national life expectancy estimates.[6] All presented figures are shown in PPP-adjusted 2019 USD (exchange rates accessed through World Bank Open Data, World Bank, n.d.).

2.3. *Policy choice experiment design*

In addition to measuring VSL, we conducted two PCEs measuring the ethical preferences between different types of outcomes for their community. Both ask the respondent to take the perspective of a community decision-maker, similar to the stance used by studies taking a societal preference approach to measure health and non-health benefits or prospective treatment (such as Kwon *et al.,* 2017; Sheen *et al.,* 2023).

In the first, we presented a trade-off between two hypothetical health interventions that save lives of people in different age groups at different levels in the respondent's community. This method followed closely the protocol used by Johansson-Stenman *et al.* (2011) to estimate the relative value of lives of different age groups, without dependence on WTP methodologies.

Next, we designed a choice experiment to further assess the relative value of life-saving and income-increasing interventions. Interviewers described the concept of cash transfers and asked respondents to actively think of and list the potential benefits of cash transfers to a household in their community. Respondents were then presented with a series of choices for how to allocate a hypothetical set of development programs in their community, which saves lives in expectation and allocates cash transfers to households in their village. We vary the number of cash transfers and the number of children saved across the different choices, in order to extract an estimate for the value of cash transfers in terms of lives saved. The design of both choice experiments is summarized below in Table 1. To aid comprehension, we used visual aids displaying each choice for both choice experiments to aid comprehension.

To check comprehension, in both choice experiments, the respondent was first presented with a dominance test, in which they were presented with a choice where one option was strictly superior (e.g., a program that saves 100 lives of people aged 19–40 vs. a program that saves 500 lives of people aged 19–40). A consistency test was also included in each choice set, in which the same choice was presented later in the survey to see if they made the same choice again. In our final analysis, only respondents who passed these comprehension tests were included.

Analysis for the choice experiments was conducted by examining the distribution of choices and by using a logit model to estimate the relative values placed on different outcomes. The estimation assumes that respondents have homogeneous preferences and that the utility of the choice is linear in the number of lives saved as well as the number of

[6] We expect this value to be approximate only, as given the poverty level in our sample, the life expectancy of our sample may vary considerably from the national life expectancy of Kenya/Ghana, and individuals may also not have an accurate estimate of their own longevity. Therefore, we focus on the range of possible VSLY and directional comparisons to existing thresholds and prior studies in our interpretation of the results (rather than precise estimates).

Table 1. *Summary of policy choice experiment design*

Attribute	Levels	Choice set
Policy choice experiment A: Life–saving, different age groups		
Number of people saved	100/200/300/400/500	Participants are presented with three choices where the order and values presented randomized.
Age of people saved	Under 5/5–18/19–40/over 40	
Policy choice experiment B: Life–saving and cash transfers, at different relative rates		
Number of children under 5 saved	5/6[a]	Participants are presented with three choices where the order and values presented were randomized. If the respondent did not switch between program A/B during presented choices, extreme options were presented (1,000, then 10,000 cash transfers for Program B).
Number of households receiving a $100 cash transfers	15/25/35/45/55/65/75/85/95/105	

[a]During the piloting, we found that this policy experiment worked best when the differential of numbers of lives saved was limited. Based on our qualitative work, increasing the number of levels led to more emotional responses limiting respondents' engagement and acceptance with the scenarios. Age was not varied, and we focused on children under 5 given the greatest policy relevance of this age group to the decision-maker, GiveWell.

cash transfers given. We take the ratio of the coefficients on the difference in the number of lives saved to that in the number of cash transfers given to be the relative value (where standard errors are calculated using the delta method). In these estimations, we only use results from the respondents' first three choices, since the latter choices were only presented to respondents who did not switch earlier and were hence not independent. The model used is given in full in Appendix B.

3. Results

3.1. Demographics

A summary of demographics for our sample can be found in Table 2. Our sample was 54 % female, with an average age of 40.39, and 44 % literate (can both read and write). Our approach identified a sample of households living in extreme poverty, with an estimated consumption per capita significantly below the national averages for both Kenya and Ghana (sample in Kenya $310.91 vs. Kenya 2019 national average $1970.11; sample in Ghana $428.49 vs. Ghana 2019 national average $2167.91 – note these consumption numbers are raw, PPP-adjusted consumption is presented below in Table 2).

3.2. Small probability comprehension (VSL)

Following our small probability training module, 58 % of our sample answered all four basic probability and risk reduction test questions correctly the first time (see Appendix A for

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

Table 2. *Summary of sample demographics*

Variable	Aggregate: Mean/%	Kenya: Mean/%	Ghana: Mean/%
N	1,820	905	915
Female	54 %	54 %	54 %
Age	40.4 (range 18–88)	42.3 (range 18–88)	38.49 (range 18–80)
Literacy: can read and write	44 %	68 %	19 %
Christian	61 %	79 %	42 %
Muslim	29 %	12 %	47 %
Household size	8.5 (SD: 5.9)	6.39 (SD: 3.11)	10.6 (7.1)
Number of children in the household	4.4 (SD: 3.6)	3.4 (SD: 2.5)	5.3 (4.2)
Has long–term health condition	35 %	41 %	29 %
Has received any cash transfer from government or charity	16 %	10 %	23 %
Has received any other charity assistance	35 %	50 %	20 %
Urban	21 %	20 %	22 %
Annual consumption per capita (PPP 2019 USD[a])	750.9 (816.4)	757.1 (734.3)	744.7 (897.6)

[a]Conversion rates sought from https://data.worldbank.org/ (World Bank, n.d.), all consumption data presented are 1 % winsorized to account for a small number of unreasonably high results in our consumption module that we expect were due to data entry error.

specific questions and results). Following the literature (e.g., Hoffmann *et al.,* 2012) we construct subsamples that exhibit "sufficient" comprehension by dropping respondents from our main VSL estimates who demonstrate one of four flags of poor understanding (see Table 3). Sixty-two-point-nine per cent of our sample passed these flags and are included in our core WTP and VSL estimates. Estimates based on the full sample, and samples constructed using alternate comprehension flags were also made and used for sensitivity analysis around our results (see Appendix C).

3.3. VSL results

We measured an average WTP of \$392.9 (SE: \$32.8) for the intervention offering 5 in 1,000 risk reduction over 10 years, or equivalent to a central VSL estimate of \$78,576.8. Depending on the assumptions made and subsample considered (see Table C1 in Appendix C), we found VSL range around this core estimate of \$66,795 to \$90,453.

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

Table 3. Summary of probability comprehension results for the full sample, and disaggregated by country

Flags of small probability comprehension	Full sample (n = 1,820)	Ghana (n = 905)	Kenya (n = 915)
1. Answered at least one basic comprehension question incorrectly, despite one attempt to explain (2)	20.2 %	13.6 %	26.7 %
2. Incorrectly answered mortality risk question on the first attempt (3)	16.6 %	12.9 %	20.3 %
3. Preference for the vaccine with lower risk reduction even if sold at same price (8)	7.4 %	4.8 %	10.1 %
4. WTP more for the vaccine with the lower risk reduction (7)	11.6 %	9.6 %	13.6 %
Proportion that failed at least one flag	37.1 %	27.9 %	46.3 %
Proportion with "sufficient" understanding, sample used for VSL estimates (passed all four flags)	62.9 %	72.1 %	53.7 %

Eighty-four-point-eight per cent of the households included in our sample contained children under the age of 18, and 74.5 % of our sample respondents were parents to at least one of the children in the house. Of the children randomly selected for the child WTP questions 44.9 % were aged 0–4 years ("under 5") and 55.1 % were aged 5–17 years ("5 and over"). We present our results divided into these two age groups for best comparison with target development programs (many of which target children under 5 specifically), for best comparison with the age group choice experiment results below, and to group observations into two groups for greater estimate precision.

We measured an average WTP for the 5 in 1,000 risk reduction children under 5 of \$465.1 (SE: \$67.3) and of 5 and over of \$382.8 (SE: \$35.8) (Table 4).

We found evidence of internal and external validity of our VSL results. Across our full sample, 70.0 % of respondents strictly passed the internal scope test (i.e., were willing to pay strictly more for the higher risk reduction). Taking our main sample of respondents with sufficient comprehension, the average WTP for low- versus high-risk intervention showed a significant difference of \$361.4 (SE: 19.4, $p < 0.01$ on t-test) across all responses. Those who were shown the intervention offering a 10 in 1,000 risk reduction first were willing to pay \$215.7 more (SE: 55.6, $p < 0.01$ on t-test) than those randomly offered a 5 in 1,000 risk reduction. In other words, our sample passes the weak external scope test (there is sensitivity to size of risk across our population) but fails the strong external scope test (that sensitivity was not proportionate to the size of the risk). We found a positive correlation between WTP and annual household consumption, reported life satisfaction, male respondent and region (full regression of respondent characteristics on WTP can be found in Appendix C).

3.4. Policy choice experiment

We found good evidence of respondent comprehension of our PCEs. First, considering choice experiment A, which compared live saving among different age groups, 97.1 % chose

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

Table 4. *WTP results and estimated VSL for both adults, and their children, for the sample of respondents that demonstrated sufficient understanding, and internal and external scope test results*

	WTP results			VSL results		
	N	WTP PPP USD: Mean	WTP USD: SE	Estimated VSL (full sample), PPP USD	Estimated VSL (Ghana sample[a]), PPP USD	Estimate VSL (Kenya sample[a]), PPP USD
Adult WTP results, subsample with "sufficient" comprehension						
WTP for 5 in 1,000, all responses[b]	1,126	392.9	32.8	78,576.8	71,566.8	88,079.9
WTP for 10 in 1,000, all responses	1,126	754.3	54.0	75,426.2	74,605.7	76,538.5
WTP for 5 in 1,000, presented first	544	452.3	51.8	90,453.5	81,583.0	101,119.7
WTP for 10 in 1,000, presented first	582	668.0	68.0	66,795.3	68,037.4	64,908.0
Child under 5 WTP results, subsample with "sufficient" comprehension						
WTP for 5 in 1,000, all responses	447	465.1	67.3	93,029.1	74,735.2	121,462.9
WTP for 10 in 1,000, all responses	447	854.2	113.4	85,424.4	79,918.2	93,982.7
Ratio child under 5 to adult (based on main estimate: WTP 5 in 1,000, all responses with sufficient comprehension)				1.2	1.0	1.4
Child 5–18 WTP results, subsample with "sufficient" comprehension						
WTP for 5 in 1,000, all responses	522	382.8	35.8	76,561.9	128,972.3	69,028.4
WTP for 10 in 1,000, all responses	522	660.4	77.0	66,035.8	75,301.0	56,108.7
Ratio child 5–18 to adult (based on main estimate: WTP 5 in 1,000, all responses with sufficient comprehension)				1.0	1.8	0.8

[a]Note that the presented VSL was obtained from a sample of lower-income households, and so do not represent VSL for the whole of Kenya/Ghana.
[b]This estimate is considered the main central estimate for VSL in the remainder of this paper.

Table 5. *Choice experiment results for full sample, and disaggregated by country*

	Full sample	Kenya	Ghana
Policy choice experiment A: Life–saving, different age groups			
N (passed both comprehension tests)	1,493	703	790
Under 5 (reference age group for analysis)	1	1	1
Relative value of individuals aged 5–18 (SE)	0.79 (SE: 0.002)	0.90 (SE: 0.003)	0.64 (SE: 0.008)
Relative value of individuals aged 19–40 (SE)	0.11 (SE: 0.002)	0.20 (SE: 0.002)	−0.03 (SE:0.009)
Relative value of individuals aged over 40 (SE)	−0.71 (SE: 0.012)	−0.27 (SE: 0.006)	−1.52 (SE: 0.131)
Policy choice experiment B: Life–saving and cash transfers, at different relative rates			
N (passed both comprehension tests)	1,601	794	807
Logit model central estimate (SE) – including those not willing to switch[a]	91.1 (SE: 44.2)	14.5(SE: 67.2)	200.9 (SE: 1,352.4)
N (%) not willing to switch from intervention saving more lives at any value	685 (37.8 %)	217 (23.8 %)	468 (52.1 %)

[a]We only use results from the respondents' first three choices, since the latter choices were only presented to respondents who did not switch earlier and were hence not independent. We conducted sensitivity analysis to these results by removing these observations entirely from the sample, but for reasons discussed below ultimately chose to include them. As a result of including them, we accept an increase in the SE of the estimates.

the strictly superior program in our dominance test, and 84.8 % chose the same program when the same choice was presented again later in the survey, demonstrating consistency. Eighty-two-point-three per cent passed both tests and so were included in the analysis. For choice experiment B, which compares saving lives to cash transfers, 95.3 % passed the dominance test and 92.1 % passed the consistency test. Eighty-eight-point-four per cent passed both tests and so were included in the analysis.

In choice experiment A, we found the highest valuation of the youngest age group, children under 5, with relative valuation decreasing with age. We found consistently low relative valuation of lives over 40 compared to all presented age groups (further interpretation of the negative valuation is in Section 4).

From choice experiment B, our central estimate of the relative value of live saved relative to a number of $1,000 cash transfers using a logit model was 91.1 (SE: 44.19) (Table 5). In other words, saving the life of a single child under 5 in the respondents' communities is equivalent to giving $1,000 cash transfers (nominal) to approximately 91 poor households (or a nominal derived value per life saved of $91,100). This result, however, is significantly skewed by the large number of individuals in our sample who were never willing to switch to a program that saves less lives regardless of the number of cash transfers offered in the

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

alternate option. It also results in very large variation in results between Kenya and Ghana given the much higher proportion of individuals unwilling to switch at any value in our Ghanaian sample.

4. Discussion

4.1. VSL in low-income settings

From our main estimate model, we find adult VSL for the study population ranging from \$66,795.3 to \$90,453.5 and VSLY ranging from 3.9 to 5.3 times the estimated average annual consumption of our participants.[7]

Our results show some sensitivity to the analysis approach (e.g., how tightly the sample is restricted to individuals with the best understanding of small probabilities and the order in which risk levels are presented). Yet overall, our estimation of VSL is relatively consistent across measurement approaches. We also find reasonably high levels of comprehension of our methods, as illustrated through passing of both the internal and weak external scope test, and reasonable levels of small probability comprehension after training, comparable to similar studies, for example, among an urban, poor population in Bangladesh (Mahmud *et al.,* 2011).[8]

4.2. Evaluating the current "benefit transfer" approach

The captured VSL results are higher than those predicted by the "benefit transfer" approach, as described by Robinson *et al.* (2019), which uses GDP per capita differences to "transfer" VSL estimates from HICs to LMICs. While our range overlaps with the upper end of estimates of VSL for a population with the same income level as our sample (following the reference case approach), our central estimate is over double that predicted from the preferred option for VSL estimation (which assumes an income elasticity of 1.5).[9] Similarly, our full range of estimates for VSLY (3.9–5.3) are consistently higher than the current suggested cost-effectiveness threshold of 1–3 times GDP per capita.

These results add to a growing number of studies suggesting that current benefit transfer approaches and cost-effectiveness thresholds based on extrapolation from HIC data using relative income may be underestimating the value individuals in LMICs place on health

[7] Since our study population was largely rural and engaged in home production of food, we used annual per capita consumption to proxy income. The resulting rations are higher than the WHO guidance of between 1 and 3 times per capita GDP for "cost-effective" health interventions. Given our study population likely has a lower life expectancy than the general population of Kenya/Ghana, we expect this to be a conservative estimate. We also note that our sample was designed to provide a reasonable representation of the intended recipients of GiveWell "top charities." As such, we targeted specific regions of Kenya and Ghana, and only poor households within each selected community. Therefore, our results are not representative of either country.

[8] We included two directly comparable questions. In a sample of respondents in urban Bangladesh, Mahmud *et al.* found that 74 % and 83 % answered a question about two lotteries and two roads, respectively. In our sample, 81 % and 93 % answered the same questions correctly.

[9] Based on an assumed income elasticity of 1.5, extrapolating from a USA VSL of \$9.4 million, and relative USA GNI per capita of \$57,878 – following the preferred option laid out in Robinson *et al.* (2019) – we estimate a VSL for this population of approximately \$14,000. Based on the alternate transfer options provided by Robinson *et al.*, we find an estimated VSL of \$76,000 (100 times GNI per capita), and \$121,000 (160 times GNI per capita).

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

interventions. Using a similar WTP approach, Trautmann *et al.* (2021) in Burkina Faso and Patenaude *et al.* (2019) in Tanzania found similar results, with comparable VSLY in the range of 3.5 to 6 times GDP per capita.

If further corroborated, a higher VSL in lower-income settings would mean that current thresholds are undervaluing the benefits of interventions that avert mortality. This in turn could mean that a substantial number of health interventions previously deemed to fall below thresholds, should be considered cost-effective. Given the significant policy implications of this finding, there is a clear need to continue to conduct further VSL research in LMICs building up a stronger set of data points, across more settings.

4.3. Value of risk reduction for child fatality

Consistent with research in HICs, our VSL elicitation found that WTP for risk reduction decreased with age and overall, child VSL was higher than adult VSL for this population. WTP was highest on average for children under 5 (1.2 times the average WTP for adults). However, we also note that our results fall well below the 2.0 average ratio typically found in HICs and that our WTP for children aged 5–18 was not significantly different from the average WTP for adults.[10] This lower estimate is consistent with the only other study of a comparable population, a pilot study by Odihi *et al.* (2021) which found a 1.0 estimated ratio of child to adult VSL. However, with only two relatively small sample studies, clearly more data points on child VSL in LMICs are required. Furthermore, as discussed further below, using our alternate estimation approach, we found a much higher preference for avoiding child fatality than through our VSL approach.

4.4. Methodological learnings from capturing VSL in low resource settings

The biggest challenge with implementing this survey in an LMIC was ensuring respondents adequately understood the questions and that we had sufficient evidence of that understanding to support data use by the decision-maker. VSL elicitation through stated preference relies on comprehension of small probabilities and mortality risks which can be especially difficult to convey in a population with limited education (43 % of our sample had not completed any primary school) and literacy (56 % were unable to read and write). We drew extensively on risk communication literature during the adaptation of the tool and piloting phase.

For future stated preference studies in LMICs, we would recommend (i) an extensive training module in probability concepts that started with easy and tangible examples (almost all respondents had some preexisting sense of chance via a lottery or coin toss, so we started there), (ii) the use of visual aids (we used standard risk grids to communicate risk out of 1,000 and created our own visual aids to convey with symbols the different aspects of choice experiments), (iii) repiloting the survey in every new location and language, and (iv) include multiple varied questions/assessments of understanding.[11]

As noted above, with these efforts, we obtained levels of comprehension comparable to those obtained in samples with a higher proportion of literacy/education. However, many

[10] Although we note a large country-level variation and wide SE for these results given the smaller sample size that had children in this age group in each country.

[11] For example, including test questions, recording closely the number of additional explanations required and recording a subjective judgment of overall respondent understanding by the surveyor.

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

respondents were unable to correctly answer questions related to the scale of risk (i.e., we found good comprehension of the difference between 5 in 1,000 and 10 in 1,000, but poor comprehension of the difference between 10 in 1,000 and 10 in 10,000). This is important given that extrapolation of VSL from stated preference relies on the assumption that WTP is proportional to risk level. However, denominator neglect and poor comprehension of risk scale have been well documented across all settings, even in populations with high literacy (Garcia-Retamero *et al.*, 2012). As such, we see this as a drawback of VSL overall, as opposed to a particular challenge with its application in low-income populations.

Contemporaneous qualitative work alongside the stated preference study proved helpful to lend credibility to findings, especially as comprehension was expected to be a challenge. Brief qualitative questions for a random proportion of respondents, and in-depth qualitative interviews with a smaller sample allowed us to understand the decision process and rationale behind the preferences expressed by our respondents. For example, we were able to support our finding that life-saving interventions for children were valued higher than for adults by demonstrating that respondents engaged with the scenario and gave clear ethical rationales similar to those found in HICs (such as the fair innings argument, the importance of protecting the weakest/ voiceless and to preserve the potential future economic and emotional value of the child).

4.5. Comparing the policy choice experiment with VSL

We compare the PCE to traditional VSL elicitation on two dimensions: the success of field elicitation and the implied VSL estimated from each approach.

As hypothesized, our results suggest that the PCE was easier to understand for respondents. In the traditional VSL approach, only 62.9 % of respondents passed all of our tests for small probability comprehension and were therefore used in the estimation. For the PCE experiments, 82.3 % (PCE A) and 88.4 % (PCE B) of the respondents passed our comprehension tests and were therefore included. However, the PCE experiment also had a clear downside in that 37.8 % of households did not switch away from saving lives for any amount of cash transfers. This problem was especially acute in Ghana, where 52.1 % of respondents never switched.

Although we are able to include these households in the estimation, we consider estimation including these households as less reliable. Our qualitative work identified that a proportion of the people unwilling to switch away from the life-saving intervention at any extreme did so because of skepticism about how cash would be disbursed or used (i.e., not focused on the real value of the outcomes). Such considerations make the comparison with cash transfers less useful for an estimate of VSL. Additionally, the quantitative estimation model is less reliable when a large proportion of respondents do not switch, because switching is necessary to establish bounds on the value of life from any individual respondent. Overall, we find that both the VSL and PCE elicitation approaches were feasible in our low-income sample, but both came with significant challenges.

Turning to comparison of estimation of VSL, the results of the PCE were directionally consistent with traditional VSL elicitation, but with some key differences (Table 6). For adults under 40, the PCE gives a VSL that is roughly one third of the traditional elicitation approach ($25,000 vs. $83,715). However, the PCE gives a much higher estimate of VSL for children compared to the traditional approach at 2.45 times for children under 5 and 2.36 times for children aged 5–18.

Table 6. *Comparing VSL and policy choice experiment results*

Age group	VSL				PCE		
Age	Ratio[a]	VSL[b]	Bounds (SE)	Bounds (choice of estimate)	Ratio[a]	Implied Value[b]	Bounds (SE)
Child under 5	1.11	$93,029	$82,918–$103,139	$65,552–$102,920	9.09	$228,755	$116,908–$345,663
Child 5–18	0.91	$76,562	$69,525–$83,597	$64,685–$107,415	7.18	$180,716	$92,357–$273,074
Adult under 40	1.00	$83,715	$77,774–$89,705		1.00	$25,163	$12,859–$38,023
Adult 40 and over	0.86	$71,744	$64,457–$79,030	$66,795–$90,453	−6.45	−$162,416	−$245,421 to −$83,005

[a]Adults under 40 are used as the reference point for both ratios.
[b]Purchasing power parity conversion applied for all values.

Further than that, the PCE experiment actually estimates a negative value of lives for individuals over 40. This stems from the fact that many of our respondents never switched when given a trade-off between saving lives of younger community members versus older ones. In our qualitative work, we found extensive qualitative evidence to support a strong valuation of young lives, but none to support a true negative valuation of older lives.[12] A previous study in Bangladesh (Johansson-Stenman *et al.,* 2011) also estimated negative values for older community members, but the authors interpreted the results as showing a strong preference for saving life-years as opposed to saving any lives. Overall, we think that these negative results should not be taken literally and can instead be considered additional evidence that individuals place a much higher value on the lives of younger over older community members.

Taking the PCE results, our qualitative findings, and similar findings in other literature, we conclude that the strong preference for saving younger lives is likely a "true" preference that is picked up by the PCE approach but not VSL, pointing toward a benefit of the PCE approach. More research studies in low-income countries would be very helpful to verify these findings and further explore if an altered PCE design could address some of these limitations.

4.6. Toward a conceptual and practical framework

Benefit–cost analysis is founded in welfare economic theory. As such, it assumes that an individual is the best judge of their own welfare ("consumer sovereignty") and that any resulting policy decision should try to maximize that individual welfare. Individual stated or revealed preferences studies capture trade-offs between their own consumption and a good or service from which we can derive an assumed value of that good or service. VSL specifically represents an individual's willingness to trade between personal consumption/spending and a small reduction in his or her own risk of death (the individuals' marginal rate of substitution between money and fatality risk).

Relying only on BCA, and VSL as the conversion metric, for decision-making relies on two main assumptions: (i) the ultimate goal is to maximize only for individual utility (represented by optimizing for their individual preferences) and (ii) empirically captured preferences truly capture the individual's best judgment of their welfare.

However, the direct choices that decision-makers in governments or donors face are generally very different from that faced by individuals. There are other inputs critical to a final policy choice that are inherently extrinsic to personal preference, such as concerns about fairness, customs, distribution and so forth. By giving an individual a set of decisions from the perspective of a decision–maker, they are asked to incorporate their views on these factors beyond their individual preference into their choice.

Additionally, given all the methodological challenges with capturing preferences, and scarcity of relevant data in many settings, it is reasonable that decision-makers question whether available VSL estimates truly capture the individual preferences of their population.

[12] As noted earlier in the discussion, our qualitative work identified many respondents gave strong arguments pro-saving young lives (including the "fair-innings" argument, the need to protect the weakest/voiceless and the greater long-term economic potential of younger lives). However, no respondents made arguments that were explicitly anti-older lives (e.g., a negative value could be consistent with a view that there are too many older people in society, but this did not emerge as a theme in any of our work).

Presenting choice experiments that are relatively easier to grasp, closer to the decision at hand, not reliant on comprehension of small probabilities and not limited by an individual's own liquidity may feel preferable to decision-makers.

At some stage, decision-makers have a normative judgment about what factors are important to inform their decisions. If they care seriously about incorporating preferences, then they have a choice to either (i) incorporate VSL and adjust for other factors with distribution or equity weights post hoc, (ii) incorporate preferences expressed from the societal perspective where an individual has already reflected these weights or (iii) a combination of the above. Combining these perspectives does not reconcile with a single economic theory, but it does give a practical way to move forward with different imperfect data sources.

Ideally, a decision-maker would have data points from multiple sources to help them incorporate these various parameters. In our experiment, we indeed had evidence from a number of elicitation approaches and took steps to provide practical guidance as to how to combine our results to arrive at an overall, decision-relevant result. First, we triangulated our estimates across multiple methodologies. Second, we conducted an extensive sensitivity analysis within each method and of our triangulation approaches. Third, as part of the sensitivity analysis, we included the option for the user (in this case GiveWell's staff) to place their own subjective weights on the different study methods (i.e., how much weight to give to VSL estimates vs. the PCE estimates) and existing literature results (i.e., how much weight to give to estimates from this study vs. estimates obtained through the benefit transfer approach). Additional detail on this process is given in Appendix D.

5. Conclusion

Our results add to a small but growing number of studies attempting to measure WTP for risk reduction through stated preference in SSA populations living in extreme poverty. Through this approach, we have found estimated VSL higher than predicted for a population of this income level. This adds to the literature supporting that the existing approaches relying on extrapolation from HICs may be underestimating the relative cost-effectiveness or cost–benefit of health promoting programs.

Additionally, we add results from a novel "PCE," designed to overcome some drawbacks of traditional VSL elicitation when considering specific policy choices. We found similar results in the PCE versus traditional VSL elicitation, but the PCE uncovered much stronger preferences for saving lives of young versus older community members. This provides additional support to findings in the literature that suggest that policymakers consider life-years saved as opposed to a single value of VSL for the entire population.

Our results also highlight the value of conducting multiple estimation approaches with the same respondents to allow for better understanding of the underlying preferences of those individuals. This approach also allows for the triangulation of multiple data points and sensitivity analysis during which subjective judgments about which is the most useful central estimate for a benefit–cost model can be made explicit. Further work is required to both collect more data across LMICs and better understand how different methods beyond just VSL can be incorporated into other decision-making processes.

Acknowledgments. This paper builds on a 2022 conference held by the Brocher Foundation, "Healthy, Wealthy, and Wise – The Ethics of Health Valuation," organized by Nir Eyal (Rutgers University), Samia Hurst (University of Geneva), Lisa A. Robinson (Harvard University), and Daniel Wikler (Harvard University). We thank the Brocher Foundation and conference participants for their support and helpful comments. This special issue was supported by the Brocher Foundation, with supplemental funding from the Rutgers University Center for Population-Level Bioethics, the University of Geneva Institute for Ethics, History, and the Humanities, and the University of Bergen Centre for Ethics and Priority Setting. More information on the Brocher Foundation is available here: https://fondation-brocher.ch/. We would also like to acknowledge and thank GiveWell for their ongoing support throughout this research.

Competing interest. The authors declare no competing interests.

References

Banzhaf, H. Spencer. 2022. "The Value of Statistical Life: A Meta-Analysis of Meta-Analyses." *Journal of Benefit–Cost Analysis*, *13* (2): 182–197. https://doi.org/10.1017/bca.2022.9

Benkhalifa, Abdelaziz, Paul Lanoie, and Mohamed Ayadi. 2012. "Estimated Hedonic Wage Function and Value of Life in an African Country." MPRA Paper, February 2, 2012. https://mpra.ub.uni-muenchen.de/75109/.

Cairns, Matthew, Serign Jawo Ceesay, Issaka Sagara, Issaka Zongo, Hamit Kessely, Kadidja Gamougam, Abdoulaye Diallo, *et al.* 2021. "Effectiveness of Seasonal Malaria Chemoprevention (SMC) Treatments When SMC Is Implemented at Scale: Case–Control Studies in 5 Countries." *PLoS Medicine*, *18* (9): e1003727. https://doi.org/10.1371/journal.pmed.1003727

Corso, Phaedra S., James K. Hammitt, and John D. Graham. 2001. "Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation." *Journal of Risk and Uncertainty*, *23* (2): 165–184. https://doi.org/10.1023/A:1011184119153

Doty, A J. 2014. "Participatory Wealth Rankings as a Tool for Targeting and Evaluation: Do Participatory Methods Successfully Identify the Poor and Measure Change in Their Lives?" https://villageenterprise.org/wp-content/uploads/2015/12/Participatory-Wealth-Ranking-and-PPIs_November_2015_AJ-white-paper.pdf

Garcia-Retamero, Rocio, Yasmina Okan, and Edward T. Cokely. 2012. "Using Visual Aids to Improve Communication of Risks about Health: A Review." *The Scientific World Journal 2012* (May): e562637. https://doi.org/10.1100/2012/562637

GiveWell. n.d. "About GiveWell | GiveWell," August 31, 2023. https://www.givewell.org/about.

Hammitt, James K., and Lisa A. Robinson. 2011. "The Income Elasticity of the Value per Statistical Life: Transferring Estimates between High and Low Income Populations." *Journal of Benefit–Cost Analysis*, *2* (1): 1–29. https://doi.org/10.2202/2152-2812.1009

Haushofer, Johannes, and Jeremy Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *The Quarterly Journal of Economics*, *131* (4): 1973–2042. https://doi.org/10.1093/qje/qjw025

Hoffmann, Sandra, Alan Krupnick, and Ping Qin. 2017. "Building a Set of Internationally Comparable Value of Statistical Life Studies: Estimates of Chinese Willingness to Pay to Reduce Mortality Risk." *Journal of Benefit–Cost Analysis*, *8* (2): 251–289. https://doi.org/10.1017/bca.2017.16

Hoffmann, Sandra, Ping Qin, Alan Krupnick, Burmaajav Badrakh, Suvd Batbaatar, Enkhjargal Altangerel, and Lodoysamba Sereeter. 2012. "The Willingness to Pay for Mortality Risk Reductions in Mongolia." *Resource and Energy Economics*, *34* (4): 493–513. https://doi.org/10.1016/j.reseneeco.2012.04.005

IDinsight. 2019. "Measuring People's Preferences," December 28, 2023. https://www.idinsight.org/publication/measuring-peoples-preferences/.

Johansson-Stenman, Olof, Minhaj Mahmud, and Peter Martinsson. 2011. "Saving Lives versus Life-Years in Rural Bangladesh: An Ethical Preferences Approach." *Health Economics*, *20* (6): 723–736. https://doi.org/10.1002/hec.1627

Kazibwe, Joseph, Adrian Gheorghe, David Wilson, Francis Ruiz, Kalipso Chalkidou, and Y-Ling Chi. 2022. "The Use of Cost-Effectiveness Thresholds for Evaluating Health Interventions in Low- and Middle-Income Countries from 2015 to 2020: A Review. *Value in Health*, *25* (3): 385–389. https://doi.org/10.1016/j.jval.2021.08.014

Kwon, Sun-Hong, Sun-Kyeong Park, Ji-Hye Byun, and Eui-Kyung Lee. 2017. "Eliciting Societal Preferences of Reimbursement Decision Criteria for Anti-Cancer Drugs in South Korea." *Expert Review of Pharmacoeconomics & Outcomes Research*, *17* (4): 411–419. https://doi.org/10.1080/14737167.2017.1277144

Mahmud, Minhaj. 2011 "On the Contingent Valuation of Mortality Risk Reduction in Developing Countries." *Applied Economics*, *41* (2): 171–181. https://www.tandfonline.com/doi/abs/10.1080/00036840600994252

McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In Zaraembka, P. (Ed.) *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.
Mofadal, Adam I. A., Kunnawee Kanitpong, and Piyapong Jiwattanakulpaisarn. 2015. "Analysis of Pedestrian Accident Costs in Sudan Using the Willingness-to-Pay Method." *Accident Analysis & Prevention*, *78* (May): 201–211. https://doi.org/10.1016/j.aap.2015.02.022
Odihi, Deborah, Gatien De Broucker, Zahid Hasan, Sayem Ahmed, Dagna Constenla, Jasim Uddin, and Bryan Patenaude. 2021. "Contingent Valuation: A Pilot Study for Eliciting Willingness to Pay for a Reduction in Mortality from Vaccine-Preventable Illnesses for Children and Adults in Bangladesh." *Value in Health Regional Issues*, *24* (May): 67–76. https://doi.org/10.1016/j.vhri.2020.10.004
Palanca-Tan, Rosalina. 2013. "Age Preferences for Life-Saving Programs: Using Choice Modeling to Measure the Relative Values of Statistical Life." *The Singapore Economic Review*, *58* (2): 1350008. https://doi.org/10.1142/S0217590813500082
Patenaude, Bryan N., Innocent Semali, Japhet Killewo, and Till Bärnighausen. 2019. "The Value of a Statistical Life-Year in Sub-Saharan Africa: Evidence from a Large Population-Based Survey in Tanzania." *Value in Health Regional Issues*, *19* (September): 151–156. https://doi.org/10.1016/j.vhri.2019.07.009
Peasgood, Tessa, Martin Howell, Rakhee Raghunandan, Amber Salisbury, Marcus Sellars, Gang Chen, Joanna Coast, *et al.* 2024. "Systematic Review of the Relative Social Value of Child and Adult Health." *Systematic Review*, *42*: 177–198. https://doi.org/10.1007/s40273-023-01327-x
Pryce, Joseph, Marty Richardson, and Christian Lengeler. 2018. "Insecticide-Treated Nets for Preventing Malaria." *Cochrane Database of Systematic Reviews*, *11*: CD000363. https://doi.org/10.1002/14651858.CD000363.pub3
Radin, Mark, Marc Jeuland, Hua Wang, and Dale Whittington. 2020. "Benefit–Cost Analysis of Community-Led Total Sanitation: Incorporating Results from Recent Evaluations." *Journal of Benefit–Cost Analysis*, *11* (3): 380–417. https://doi.org/10.1017/bca.2020.6
Robinson, Lisa A., James K. Hammitt, and Lucy O'Keeffe. 2019. "Valuing Mortality Risk Reductions in Global Benefit–Cost Analysis." *Journal of Benefit–Cost Analysis*, *10* (Suppl 1): 15–50. https://doi.org/10.1017/bca.2018.26
Robinson, Lisa A., William J. Raich, James K. Hammitt, and Lucy O'Keeffe. 2019. "Valuing Children's Fatality Risk Reductions." *Journal of Benefit–Cost Analysis*, *10* (2): 156–177. https://doi.org/10.1017/bca.2019.10
Sheen, Daniel, Tessa Peasgood, and Ilias Goranitis. 2023. "Eliciting Societal Preferences for Non-Health Outcomes: A Person Trade-Off Study in the Context of Genomics." *Clinical Therapeutics*, *45* (8): 710–718. https://doi.org/10.1016/j.clinthera.2023.07.004
Syuhada, Ginanjar, Adhadian Akbar, Donny Hardiawan, Vivian Pun, Adi Darmawan, Sri Hayyu Alynda Heryati, Adiatma Yudistira Manogar Siregar, *et al.* 2023. "Impacts of Air Pollution on Health and Cost of Illness in Jakarta, Indonesia." *International Journal of Environmental Research and Public Health*, *20* (4): 2916. https://doi.org/10.3390/ijerph20042916
Trautmann, Stefan T., Yilong Xu, Christian König-Kersting, Bryan N. Patenaude, Guy Harling, Ali Sié, and Till Bärnighausen. 2021. "Value of Statistical Life Year in Extreme Poverty: A Randomized Experiment of Measurement Methods in Rural Burkina Faso." *Population Health Metrics*, *19* (1): 45. https://doi.org/10.1186/s12963-021-00275-y
Viscusi, W. Kip, and Clayton J. Masterman. 2017. "Income Elasticities and Global Values of a Statistical Life." *Journal of Benefit–Cost Analysis*, *8* (2): 226–250. https://doi.org/10.1017/bca.2017.12
Watts, Elizabeth, Joshua Mak, and Bryan Patenaude. 2022. "Benefit–Cost Ratios of Continuing Routine Immunization during the COVID-19 Pandemic in Africa." *Journal of Benefit–Cost Analysis*, *13* (1): 91–106. https://doi.org/10.1017/bca.2021.13
Wiegand, Martin. 2020. "Welfare Measurement and Poverty Targeting Based on Participatory Wealth Rankings." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.3756257
World Bank. n.d. *World Bank Open Data*. World Bank Open Data, August 31, 2023. https://data.worldbank.org.

Appendix A: VSL – Detailed survey questions

Small probability training and comprehension: Our contingent valuation survey module started with the surveyor providing a basic, scripted explanation of probabilities based on a lottery, which was followed by an understanding test question. They are introduced to the risk visual aids at this stage with an image that demonstrates the probability of a person winning a lottery. Next, respondents were introduced to the concept that much as in a lottery individuals face a different probability of winning, in life we often face scenarios that expose us to different levels of risk of injury or death. They are asked to imagine two roads, one of which the cars drive very fast and there

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

Table A1. *Small probability test questions, including the proportion of our sample who answered each question correctly the first time*

Questions	Percent answered correctly the first time (N = 1,820)
1. **Basic understanding question 1:** *Imagine two lotteries. The chance of winning in one lottery is 5 in 1,000, the chance of winning in the other lottery is 10 in 1,000. Which lottery has the larger chance of winning?*	81 %
2. **Basic understanding question 2:** *Imagine two roads that are prone to accidents. The risk of dying in an accident on the first road is 1 in 1,000, and on the second road is 3 in 1,000. Which road is riskier?*	93 %
3. **Scale understanding question:** *Now imagine two different roads. The risk of dying in an accident on the first road is 1 in 100, and on the second road is 2 in 1,000. Which road is riskier?*	84 %
4. **Basic understanding question 3:** *Imagine two people. The first person's chance of death is 5 in 1,000 in the next 10 years. The second person's chance of death is 10 in 1,000 in the next 10 years. Which person is more likely to die in the next 10 years?*	84 %
5. **Risk reduction question:** *Imagine a disease that kills 50 in 1,000 people. There are three different vaccines available for the disease. Vaccine A reduces the risk of dying from this disease from 50 in 1,000 to 20 in 10,000, Vaccine B reduces the risk from 50 in 1,000 to 40 in 1,000, and Vaccine C reduces the risk from 50 in 1,000 to 30 in 1,000. Which vaccine would you choose?*	34 %

are a lot of accidents, and on the other there are efforts to slow drivers and reduce accidents. This is followed by two understanding questions – one simple, and one capturing whether respondents understand the scale of risk (i.e., can tell the difference between 1 in 100 vs. 2 in 1,000). Next, we introduce and test understanding of the concept of mortality risk from an illness over the course of a specified time period (10 years). Finally, we introduce and test the concept that their actions you can take to reduce risk over that specified time. This sets them up to start the WTP questions in the following section.

For each question, we record the individual's first response. Then if that answer is incorrect, the surveyor is prompted to give a further scripted explanation using the visual aids and re-ask the questions. We allow the surveyor to re-explain up to twice, and record how many explanations are required before the respondent answers correctly. In Table A1, we summarize the exact framing of each test question, and the proportion that answered correctly first time.

Next, we captured respondent WTP for small risk reductions. First, the scenario is introduced, and the baseline risk set as follows:

Imagine a new disease that affects (ADULT/CHILDREN) in your village. The disease is rare, so there is not much chance of you catching the disease. For every 1,000 people, 20 will catch the disease in the next 10 years. However, everyone who catches the disease will die. So, your risk of dying from the disease is 20 in 1,000 over the next 10 years (or 20 in 10,000 each year).

Next, we introduce the vaccine/medicine (randomized in Kenya, medicine only in Ghana):

A new (VACCINE/MEDICINE) has been made for the disease. It reduces your risk of dying from the disease from 20 in 1,000 to (15 in 1,000/10 in 1,000) over the next 10 years (or 20 in 10,000 each year to (15 in 10,000/10 in 10,000)). However, it is not available at the public health facility so you must buy it for yourself.

Next, we capture the initial WTP, using a payment card to help respondents come up with a figure.

Do you want to buy this (VACCINE/MEDICINE) for (YOURSELF/YOUR CHILD)? If yes, how much would you be prepared to spend today to buy this (VACCINE/MEDICINE)?

Next, we capture WTP in small installments for the vaccine/medicine over the course of 10 years (the same time frame as the risk reduction). We allow the respondent to define how frequently they want to make payments.

Now imagine that you are able to pay for this (VACCINE/MEDICINE) little by little over the next 10 years. How much are you willing-to-pay each month/year/total to receive this (VACCINE/MEDICINE), and so reduce your risk for the next 10 years?

We now repeat for a new vaccine/medicine with the higher/lower risk level.

A second new (VACCINE/MEDICINE) has been made for the disease. It reduces your risk of dying from the disease from 20 in 1,000 to (15 in 1,000/10 in 1,000) over the next 10 years (or 20 in 10,000 each year to (15 in 10,000/10 in 10,000)). However, it is not available at the public health facility, so you must buy it for yourself.

Finally, as an additional understanding check, we asked the respondent to choose directly between the two vaccines that offer different risk reductions:

If they were the same price, would you prefer to buy the first or second (VACCINE/MEDICINE).

There were a number of randomized components in this section:

(ADULT/CHILD): Respondents who were main caretakers of children under 18 in the household were asked for their WTP for *both* themselves and for a randomly selected child. Respondents who were not main caretakers of any children under 18 in the household were only asked for WTP for themselves. The order in which the questions relating to adult/child appeared was randomized to account for any ordering effect.

(VACCINE/MEDICINE): The risk reducing item offered to the respondent was randomized such that half were asked WTP for a vaccine only, and the other half were asked WTP for a medicine only. This aimed to test for sensitivity to framing.

(15 in 1,000/10 in 1,000): Respondents were offered two vaccines, offering either 5 in 1,000 and 10 in 1,000 risk reduction over 10 years. The order in which these vaccines were presented was randomized such that half received 5 in 1,000 first, and half received 10 in 1,000 first. This allowed us to test for (i) population-level scope sensitivity in WTP, by testing if respondents were on average WTP more for a higher-risk reduction (external scope test) and (ii) individual-level scope sensitivity in WTP, by testing if respondents offered more for the more effective vaccine (internal scope test).

Appendix B: Policy choice experiment analytical model

We assume uniform preferences among all individuals. (This is an unrealistic assumption. However, we do not have enough choice data from each individual to estimate their preferences. Furthermore, we care about aggregated preferences, which is a natural interpretation of what we are estimating.)

An individual is choosing between two options, i and j, which are combinations of a number of lives saved and a number of cash transfers (of 1,000 USD) given to extremely poor households. Suppose the utility that an individual derives from an option is a linear function on the number of lives saved and the number of cash transfers given, with an error term:

$$u_i = \beta_1 X_{1,i} + \beta_2 X_{2,i} + \epsilon_i,$$

$$u_j = \beta_1 X_{1,j} + \beta_2 X_{2,j} + \epsilon_j.$$

We assume that the error terms, ϵ_i and ϵ_j, are independently and identically distributed, following an extreme value distribution.[13] β_1 and β_2 represent the values that the individual places on the number of lives saved and cash transfers given, respectively, and their ratio the relative value between the two.

Option i would be chosen over option j if and only if it is associated with a higher level of utility. Hence, we can express the probability that option i is chosen over option j as

$$Pr\left(u_i > u_j | \vec{X}\right) = Pr\left(\epsilon_j - \epsilon_i < \beta_1 \left(X_{1,i} - X_{1,j}\right) + \beta_2 \left(X_{2,i} - X_{2,j}\right)\right),$$

where $\vec{X}$ represents the vector $\left(X_{1,i}, X_{1,j}, X_{2,i}, X_{2,j}\right)$.

Since $\epsilon_j - \epsilon_i$, the difference between two identical random variables following the extreme value distribution follows the logistic distribution, the above probability can be expressed using the cumulative distribution function of the logistic distribution:

[13] The probit model has identical assumptions, except that the error terms are independently and identically distributed following a standard normal distribution, and hence so do their differences.

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

$$Pr\left(u_i > u_j|\vec{X}\right) = \frac{1}{1+e^{-\left(\beta_1\tilde{X}_{1,i,j}+\beta_2\tilde{X}_{2,i,j}\right)}},$$

where $\tilde{X}_{1,i,j} = X_{1,i} - X_{1,j}$ and $\tilde{X}_{2,i,j} = X_{2,i} - X_{2,j}$ are the difference between the number of lives saved and that between the number of cash transfers given between the two options.

We run a logit model of the choices on the differences in numbers of lives saved and numbers of cash transfers given (McFadden, 1973), using all choices made by all individuals, in order to estimate the parameters β_1 and β_2.[14] To obtain the relative value, or the value of a life saved expressed in terms of the number of 1,000 USD cash transfers given, we obtain the ratio β_1/β_2 using the "nlcom" command in Stata which uses the delta method to computer the point estimate and standard error of the ratio.

In the case where we study the relative values individuals place on people of different age groups, the specification of the utility function of a choice i is

$$u_i = \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \beta_4 X_{4,i} + \epsilon_i,$$

where $X_{1,i}, X_{2,i}, X_{3,i}$ and $X_{4,i}$ represent the number of lives saved in each of the following age groups: under 5, 5–18, 19–40 and above 40. We focus on relative values individuals place on people in the latter three age groups relative to someone under 5: β_2/β_1, $\beta_3/$ and β_4/β_1.

Appendix C: Further WTP results

In Table C1, we present our full WTP results, including the range of estimated VSL, for the sample of respondents in Kenya and Ghana in local currencies (2019 nominal Kenyan Shilling, and 2019 nominal Ghanaian Cedi) and for the aggregate sample (2019 PPP-adjusted USD, World Bank, n.d). First, we present the full adult results using our "Model A" sample, which excludes respondents who do not pass basic comprehension tests (this is the same as the values presented in the main text of the paper). For comparison, we also present our results for the full sample of respondents. Finally, we present the WTP values for the same hypothetical medicine/vaccine for a randomly selected child of the respondent.

In Table C2, we present a regression of willingness to pay on respondent characteristics for adult VSL. We found significant associations between higher WTP and higher reported life satisfaction, male gender, non-Muslim religion and higher household consumption. There were also significant regional differences in WTP. Note that for this regression, we include our sample of "wealthy" respondents.

[14] The standard errors are clustered at the individual level.

https://doi.org/10.1017/bca.2024.10 Published online by Cambridge University Press

Table C1. *Extended WTP and VSL results for adult and children, for the full sample and disaggregated by country*

	Aggregate (PPP USD 2019)				Kenya (2019 Shilling)			Ghana (2019 Cedi)		
	N	$	*SE*	*VSL*	*N*	*Sh*	*SE*	*N*	*Ce*	*SE*
Adult Sample: Model A (passed comprehension flags, main estimation model)										
WTP for 5 in 1,000 all responses	1,126	393	33	78,577	478	18,448	1,431	648	726	65
WTP for 10 in 1,000 all responses	1,126	754	54	75,426	478	32,062	2,412	648	1,514	104
WTP for 5 in 1,000, when presented first	544	452	52	90,453	247	21,180	2,008	297	828	112
WTP for 10 in 1,000, when presented first	582	668	68	66,795	231	27,190	3,252	351	1,381	125
Adult_Full sample										
WTP for 5 in 1,000 all responses	1,791	499	35	99,869	892	20,918	1,485	899	821	68
WTP for 10 in 1,000 all responses	1,791	733	52	73,309	892	30,709	2,190	899	1,396	83
WTP for 5 in 1,000, when presented first	893	525	55	105,007	470	24,387	2,265	423	937	113
WTP for 10 in 1,000, when presented first	896	620	53	61,968	420	25,615	2,339	476	1,273	102
Child aged under 5, full sample										
WTP for 5 in 1,000 all responses	447	465	67	93,029	175	25,440	4,626	272	759	80
WTP for 10 in 1,000 all responses	447	854	113	85,424	175	39,369	6,659	272	1,622	171
Child aged 5–18, full sample										
WTP for 5 in 1,000 all responses	522	382	36	76,562	252	14,458	1,261	270	1,309	129
WTP for 10 in 1,000 all responses	522	660	77	66,036	252	23,503	2,561	270	1,529	186

Table C2. *Regression of willingness to pay on respondent characteristics for adult VSL*

Variables	Coefficient (*SE*)
Risk reduction: 5/1,000	−25.62 (17.07)
Female	−82.47*** (23.26)
Age	−1.67** (0.67)
Urban	35.51 (23.97)
Can read	1.73 (22.83)
Christian	−20.78 (51.41)
Muslim	−104.7* (59.31)
Household size	2.11 (1.89)
Self–reported long–term health condition	−21.05 (22.82)
Satisfaction ladder	6.41* (3.46)
Karaga (Ghana)	141.20** (55.87)
Jirapa (Ghana)	−56.49** (22.99)
Kilifi (Kenya)	122.50*** (33.97)
Log annual consumption per capita (nominal USD)	49.22*** (11.47)
N	2,004
R^2	0.049

Note: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Appendix D: Summary of result aggregation approach

As mentioned in the discussion section, our study resulted in values for VSL at the upper end of those predicted by extrapolation from HIC values following Robinson *et al.*'s (2019) reference case approach. We also output the results of the choice experiment, which has a higher value than VSL. In order to aggregate our findings, alongside previous literature we created a decision support model in which users could adjust weights at various stages to see how the final output (a central relative value of saving a life for a certain age group) varies if different assumptions about how to interpret and integrate results are made.

Banzhaf (2022) acknowledges that even for estimating VSL in the United States where there are substantially more data available, there are elements of judgment required in determining an ultimate VSL central estimate to use. Banzhaf (2022) applies subjective weights to make these judgments explicit, allows adjustments to these assumptions to be made and conducts sensitivity analysis. Here, we use the same principle in the LMIC context to explicitly apply weights to different analysis assumptions, different methods and our results relative to literature priors.

First, the user was invited to place a relative weight on the two broad approaches we used in our survey – the individual preference approach (our main VSL results) or our PCE. This weight incorporated two stages: (i) a relevance weighting where the user could place more weight on one approach over the other based on their assessment of its relevance to the decision faced and (ii) a confidence weighting where the user could adjust weighting based on their overall confidence in the two methodologies (where the default weight was based on the proportion that passed basic understanding tests for each approach, 63 % for individual VSL, and 89 % for the PCEs). At this stage, we also gave the user the option to adjust which central estimate they used for each approach acknowledging that this selection is also prone to the analyst's judgment.

Next, we allow the user to adjust for some assumptions made in the conversion. For example, they can adjust their assumptions about the approximate income and age of the population served. Then we provide the user with the best estimates of VSL for the population based on only extrapolation from the literature (applying Robinson *et al.*'s 2019 approach). Here, the user can adjust which option is the preferred to make the extrapolation (income elasticity 1.5, the default vs. 100 times GNI per capita vs. 160 times GNI per capita). Finally, the user places weights on the results of this study, relative to the best estimates based on existing literature. This model then outputs a central estimate based on the user's weights. It also allows the user to adjust weights throughout to see directly how this impacts the final output.

The benefit of this study and our combination of estimates is that it allows the user to draw on multiple methodological approaches and decision frameworks to inform their modeling and allocation decisions, rather than relying on a single VSL estimate. A potential downside of this approach is its relative complexity – it requires an engaged user with a reasonable understanding of the technical considerations of different approaches. This was possible in this instance as GiveWell already has extensive experience and comfort with explicit modeling and assigning subjective weighting for different methods. Further work would be required to understand how to incorporate a similar approach into a different decision-making process.

Cite this article: Redfern, Alice, Sindy Li, Martin Gould, Felipe Acero, and Daniel Stein. 2024. "Lessons from Applying Value of Statistical Life and Alternate Methods to Benefit–Cost Analysis to Inform Development Spending." *Journal of Benefit-Cost Analysis* 15: 127–154, doi:10.1017/bca.2024.10