METHODOLOGICAL ISSUES TO CONSIDER WHEN COLLECTING DATA TO ESTIMATE POVERTY IMPACT IN ECONOMIC EVALUATIONS IN LOW-INCOME AND MIDDLE-INCOME COUNTRIES

SEDONA SWEENEY*, ANNA VASSALL, NICOLA FOSTER, VICTORIA SIMMS, PATRICK ILBOUDO, GODFATHER KIMARO, DON MUDZENGI and LORNA GUINNESS

a London School of Hygiene & Tropical Medicine, London, UK
b Centre Muraz, Bobo-Dioulasso, Burkina Faso
c University of Cape Town, Cape Town, South Africa
d National Institute for Medical Research, Dar es Salaam, Tanzania
e Aurum Institute, Johannesburg, South Africa

ABSTRACT

Out-of-pocket spending is increasingly recognized as an important barrier to accessing health care, particularly in low-income and middle-income countries (LMICs) where a large portion of health expenditure comes from out-of-pocket payments. Emerging universal healthcare policies prioritize reduction of poverty impact such as catastrophic and impoverishing healthcare expenditure. Poverty impact is therefore increasingly evaluated alongside and within economic evaluations to estimate the impact of specific health interventions on poverty. However, data collection for these metrics can be challenging in intervention-based contexts in LMICs because of study design and practical limitations. Using a set of case studies, this letter identifies methodological challenges in collecting patient cost data in LMIC contexts. These components are presented in a framework to encourage researchers to consider the implications of differing approaches in data collection and to report their approach in a standardized and transparent way. © 2016 The Authors. Health Economics published by John Wiley & Sons Ltd.

Received 14 April 2015; Revised 14 October 2015; Accepted 11 November 2015

KEY WORDS: out-of-pocket payments; catastrophic expenditures; poverty; data collection methods; economic evaluation

1. INTRODUCTION

As universal access to health care becomes a greater international priority, interest has grown in reducing the level of financial catastrophe and impoverishment caused by health-related expenditure (64th World Health Assembly, 2011). As a result, there is increased recognition that the impact of health interventions on poverty and equity should be incorporated into economic evaluations (Bill and Melinda Gates Foundation et al., 2014) – particularly in low-income and middle-income countries (LMICs) where out-of-pocket expenditures make up a large proportion of total health expenditure (World Health Organization, 2015). This is evidenced by the growing popularity of ‘extended’ economic evaluations, which incorporate assessments of the potential financial risk protection impact of an intervention or technology (Verguet et al., 2014). In the context of this growing importance of poverty impact metrics in health planning and decision making, there is need for high-quality data to estimate the impact of health expenditures on poverty and vulnerability. To date, the majority of research reporting the poverty impact of health expenditures has drawn on data from large cross-sectional surveys such as the Living Standards Measurement Survey or World Health Survey. While these datasets facilitate equity...
analyses evaluating the distribution of health impacts or financial pooling mechanisms across socioeconomic status analysis at the national level (Xu et al., 2003; Lu et al., 2009), they cannot be easily used to capture the impact of a specific health intervention on poverty and may not always include detail on indirect costs or income loss, which can be key aspects of the poverty impact of illness.

Collecting this type of data within a smaller-scale study setting can substantially increase the time and cost of data collection. Many studies therefore avoid collecting data for a poverty impact analysis altogether. Where poverty impact data are collected as part of intervention evaluations, there are notable inconsistencies in data collection methods. Systematic reviews of existing patient cost studies in LMICs highlight a lack of standard approaches across cost ingredients, data sources, sampling methodologies, and recall periods, even where the same measure of poverty impact is used (Barter et al., 2012; Tanimura et al., 2014; Kankeu et al., 2013; Alam & Mahal, 2014; McIntyre et al., 2006). This can lead to challenges in assessing the comparability, quality, and accuracy of results. In part, this heterogeneity may stem from limited practical guidance or standards on collecting patient-incurred cost data. Reporting guidelines for economic evaluations largely cover provider perspectives (Drummond & Jefferson, 1996; Husereau et al., 2013) and are neither updated to reflect information necessary for poverty impact metrics nor provide guidance when constraints in data collection require compromise, such as limiting the sample size or restricting the length of the questionnaire.

The aim of this letter is to highlight challenges faced in collecting data on patient costs within economic evaluation platforms in LMICs. We discuss practical issues around collecting patient-incurred cost and household income data, including comprehensiveness of the survey instrument, timing of interviews, sampling, and survey administration. To illustrate these issues, we use four case studies from our own research as examples (Kufa et al., 2014; Foster et al., 2015; Ilboudo et al., 2014; Mfinanga et al., 2015) (Table I). Finally, we present a framework of methodological choices in planning research on poverty impact metrics (Table II) to encourage researchers to report their approach in a standardized and transparent way and to consider potential implications of varying approaches in data collection (Figure 1).

2. COMPREHENSIVENESS OF SURVEY DESIGN

There is a rich theoretical literature on the measurement of affordability in health care. The most common indicators of poverty impact are catastrophic expenditure (defined where health spending exceeds a threshold percentage of household income) and impoverishing expenditure (defined where health spending pushes a household below the poverty line) (Wagstaff, 2011; Wagstaff & van Doorslaer, 2014; Wagstaff & Eozenou, 2003). A number of theoretical challenges are associated with estimating the poverty impact of illness, which are not addressed in detail in this letter, including the appropriate denominator and thresholds for analysis and how to represent the long-term impact of health spending (Xu et al., 2003; McIntyre et al., 2006; Russell, 1996; Flores et al., 2008; Niëns et al., 2010; Niëns & Brouwer, 2013; Pal, 2012; Wingfield et al., 2014; Onoka et al., 2011; Moreno-Serra et al., 2011; Chuma et al., 2006; Sauerborn et al., 1996; Kruk et al., 2009). The data required are defined by the metric of poverty impact chosen but can include data on direct out-of-pocket expenditures for health care, any indirect costs of time associated with being ill or accessing care, and any further economic impact measures such as income loss or loan interest.

The main challenge in survey design is the representation of complex patient experiences within a manageable survey length. Survey length is of particular concern when a patient cost questionnaire follows a lengthy clinical investigation, as it increases the risk of survey fatigue and participation refusal and increases resources required to conduct the survey. Our four case studies had a range of survey durations; this is largely a function of the complexity of the patient pathways in question. MERGE and XTEND attempted to cover the overall costs of a complex illness episode over a range of different providers, whereas ECONPOP covered only a recent hospitalization and REMSTART covered only the current visit. Survey durations for each study are detailed in Table I.
<table>
<thead>
<tr>
<th><strong>Table I. Case study characteristics</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MERGE</strong> (Kufa et al., 2014)</td>
</tr>
<tr>
<td><strong>Country</strong></td>
</tr>
<tr>
<td><strong>Aim of study</strong></td>
</tr>
<tr>
<td><strong>Study design</strong></td>
</tr>
<tr>
<td><strong>Time frame</strong></td>
</tr>
<tr>
<td><strong>Sampling for cost data</strong></td>
</tr>
<tr>
<td><strong>Location of interview</strong></td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
</tr>
<tr>
<td><strong>Subgroups (n)</strong></td>
</tr>
<tr>
<td><strong>OOP cost ingredients</strong></td>
</tr>
<tr>
<td><strong>Recall period (costs)</strong></td>
</tr>
<tr>
<td><strong>Household/individual costs</strong></td>
</tr>
<tr>
<td><strong>Average length of interview (min)</strong></td>
</tr>
<tr>
<td><strong>Diary/recall</strong></td>
</tr>
<tr>
<td><strong>Indirect cost measurement</strong></td>
</tr>
<tr>
<td><strong>Additional health services costing</strong></td>
</tr>
<tr>
<td><strong>Source of income data</strong></td>
</tr>
<tr>
<td><strong>Interviewers used</strong></td>
</tr>
<tr>
<td><strong>(Continues)</strong></td>
</tr>
</tbody>
</table>
### Table I. (Continued)

<table>
<thead>
<tr>
<th></th>
<th>MERGE (Kufa et al., 2014)</th>
<th>XTEND (Foster et al., 2015)</th>
<th>ECONPOP (Ilboudo et al., 2014)</th>
<th>REMSTART (Mfinanga et al., 2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medium of recording</strong></td>
<td>Paper survey</td>
<td>Electronic survey</td>
<td>Paper survey</td>
<td>Paper survey</td>
</tr>
<tr>
<td><strong>Average cost (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly OOP expenditures:</td>
<td>$1.02 ($0.44 - $1.60)</td>
<td>Total OOP expenditures:</td>
<td>Total OOP expenditures:</td>
<td>OOP expenditures</td>
</tr>
<tr>
<td>Monthly travel costs:</td>
<td>$2.31 ($1.75 - $2.87)</td>
<td>Total loan interest:</td>
<td>$52.80 ($47.36–$58.24)</td>
<td>for one visit to study facility:</td>
</tr>
<tr>
<td>Monthly food costs:</td>
<td>$10.93 ($9.44 – $12.41)</td>
<td>Total income loss:</td>
<td>$54.82</td>
<td>$1.96 ($1.80–$2.13)</td>
</tr>
<tr>
<td>Monthly indirect costs:</td>
<td>$15.67 ($11.88 – $19.46)</td>
<td>Total guardian costs:</td>
<td>$32.11</td>
<td></td>
</tr>
<tr>
<td>Monthly guardian costs:</td>
<td>$6.43 ($4.59 – $8.26)</td>
<td>Total episode cost:</td>
<td>$324.07</td>
<td></td>
</tr>
<tr>
<td>Monthly carer costs:</td>
<td>$4.63 ($1.60 – $7.65)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average annual income (95% CI)</strong></td>
<td>$2,565 ($2,225 – $2,905)</td>
<td>$1,237 ($1,001 – $1,474)</td>
<td>Not measured (GDP per capita used as proxy)</td>
<td>Tanzania: $244 ($212 – 276)</td>
</tr>
<tr>
<td><strong>National poverty line (USD)</strong></td>
<td>$773</td>
<td>$773</td>
<td>$184</td>
<td>Zambia: $219 ($199–239)</td>
</tr>
<tr>
<td><strong>GDP per capita (USD)</strong></td>
<td>$6,618</td>
<td>$6,618</td>
<td>$531</td>
<td>Tanzania: $266</td>
</tr>
<tr>
<td><strong>Frequency of catastrophic expenditure at 20% threshold (95% CI)</strong></td>
<td>40% (36–45%)</td>
<td>59% (54–65%)</td>
<td>10% (6–14%)</td>
<td>4% (3–5%)</td>
</tr>
<tr>
<td><strong>Minimum sample size required to estimate frequency of catastrophic expenditure with 95% CI</strong></td>
<td>Error margin 5%: 2,282</td>
<td>Error margin 5%: 1,057</td>
<td>Error margin 5%: 13,689</td>
<td>Error margin 5%: 36,504</td>
</tr>
<tr>
<td></td>
<td>Error margin 10%: 570</td>
<td>Error margin 10%: 264</td>
<td>Error margin 10%: 3,422</td>
<td>Error margin 10%: 9,126</td>
</tr>
<tr>
<td></td>
<td>Error margin 15%: 254</td>
<td>Error margin 15%: 117</td>
<td>Error margin 15%: 1,521</td>
<td>Error margin 15%: 4,056</td>
</tr>
</tbody>
</table>

ART, antiretroviral therapy; OOP, out-of-pocket; GDP, gross domestic product; GP, general practitioner.
Disaggregation of cost ingredients will also affect survey length, and researchers may need to prioritize certain aspects to cover in depth. However, it is known that major drivers for patient costs can vary by setting and across income quintiles (Tanimura et al., 2014; Saksena et al., 2010), making it difficult to pre-suppose any exclusions or the relative attention placed on each aspect of expenditure or income measured. Surveys should be adapted to accurately represent the setting of interest, and researchers must be clear about which ingredients they do include and how ingredients are disaggregated.

Another widely recognized challenge is measurement of permanent income in LMICs, where informal employment is common and income is often seasonal (Ferguson et al., 2003; Deaton, 1997). Income data are difficult to collect in a small survey setting; as interviews in an intervention evaluation are conducted individually, accurate estimation of household income is often impossible. Researchers will need to decide whether personal income is an appropriate proxy for household income in their study context and be clear about the limitations of such a decision. In the XTEND and MERGE case studies, respondents consistently reported themselves to be the primary breadwinners in the household; personal income was therefore collected, with the limitation that these analyses may have underestimated the economic burden on the family as they did not account for the fact that income is shared amongst household members. On the other hand, within the ECONPOP sample, respondents were often not the primary breadwinners and often could not estimate household income. The decision was therefore made to use an assumption of gross domestic product per capita as a proxy rather than risk breaking the confidentiality of the interview by asking family members. This decision has implications for the metrics used; in this case, we did not have a firm understanding of where households lay in relation to the poverty line at baseline and, therefore, would not have been able to report on impoverishing expenditures.

Where researchers are unable to collect income directly, asset indices may also be used as a proxy measure of household socioeconomic position. Information on assets can be simpler to collect than income or consumption but result in ordinal data (Ferguson et al., 2003). In order to convert an asset index into monetary terms, necessary for the denominator of threshold metrics such as catastrophic or impoverishing expenditures, these data need to be mapped to an absolute wealth metric (Hruschka & Hadley, 2015; Howe et al., 2012). This may pose issues if income diversity in the population of interest is substantially different from that of the national population.

3. TIME FRAME AND RECALL

Deciding on the appropriate timing for the survey may also be difficult in a study where survey timing is based primarily on outcome measurement. The clinical pathways for some types of illness (for example, TB) can be long and complex, making recall bias a significant concern. This is illustrated in the XTEND survey, where patients enrolled in the trial could only be interviewed at the end of a 6-month follow-up period. To accommodate this, an additional sample of those on TB treatment outside the trial enrollees was also surveyed to increase sample size and allow for shorter recall periods between interviews. When capturing income loss as a result of illness in the case of complex clinical pathways, researchers will also need to weigh the risks of recall bias against the anticipated benefit of soliciting information on income before the illness.

There is also the potential for cost truncation in chronic illness or conditions with complications. The long-term economic impact of illness can be substantial (Ilboudo et al., 2013). This can be captured by following a cohort along the clinical pathway (as in the XTEND study) or with follow-up surveys conducted later (Ilboudo et al., 2013). However, it is a particular problem for lifelong treatments such as antiretroviral therapy.

Finally, dissaving or other coping strategies can also be an important reflection of the long-term impact of illness, and where possible, it may be helpful to include questions on coping strategies in the survey. Surveys may directly ask how households mobilized payment for healthcare services (Flores et al., 2008), or longitudinal surveys may be able to conduct repeated asset surveys, capturing any depletion of assets caused by illness (Ilboudo et al., 2013). This is only a partial measure of the economic impact of illness on households; however, it is a useful proxy where income measurement is impossible.
4. SAMPLE SIZE AND REPRESENTATIVENESS

Sample size considerations are key in the planning stages of a study and will depend on the aims, nature, and scope of the study, and the degree of precision (confidence interval and margin of error) deemed appropriate (Lwanga & Lemeshow, 1991). Household surveys generally follow United Nations guidelines of a 5–10% margin of error at the 95% confidence interval, with further adjustment to account for clustering and non-response (United Nations Statistical Division, 2008). However, this degree of precision may be difficult to achieve in an intervention-based context, and researchers need to be pragmatic. Some trade-off in error margin will likely need to be made in the interests of practicality of the survey; this is especially true for outcomes that are particularly rare in the population of interest, as illustrated in Table I. This decision should also be taken within the context of the larger uncertainty associated with the survey – for example, spending more time in the interview to avoid recall bias may produce more reliable results than spending additional time interviewing a great many more patients.

In each of our case studies, the sampling for out-of-pocket expenditures was restricted to a subsample of participants because of practical considerations of the study; for MERGE, XTEND, and ECONPOP, a subsample of the study population was taken, while in REMSTART, the number of follow-up visits was limited. Table I shows the sample size for each case study and the ideal sample sizes necessary for various specifications of relative precision to estimate catastrophic expenditure.

Sampling considerations pose particular issues for the estimation of impoverishing expenditures when most patients are already below the poverty line – for example, where targeting those already in poverty may be a desired feature of interventions or where investigating diseases such as HIV and TB, which disproportionately affect those below the poverty line (Bates et al., 2004). When this is the case, impoverishment becomes infrequent, making power to detect the true proportion of impoverishment very low; a different metric of poverty impact should be used in these cases. All three case studies estimating income had a large proportion of poor patients: 64% of XTEND patients, 45% of MERGE patients, and 70% of REMSTART patients had a pre-diagnosis income below the national poverty lines (Chibuye, 2014; Statistics South Africa, 2014; OECD, 2013; Laokri et al., 2013).

5. DATA SOURCES AND SURVEY ADMINISTRATION

Finally, researchers will need to identify data sources and plan administration of the survey. Researchers from the Database of Instruments for Resource Use Measurement team working in a high-income country setting (Ridyard et al., 2015) propose a taxonomy for methods of resource use measurement including the following: the source of data, who completes the resource use measurement, how it is administered, how it is recorded, and the medium of recording. Work in LMICs requires some additional consideration, as described subsequently.

Cost diaries are considered to be the gold standard in patient cost collection (Wiseman et al., 2005; Goossens et al., 2000), but they can be time and cost intensive for researchers, especially where there is high illiteracy; patient recall is more common in low-income settings (Beegle et al., 2012). This can be supplemented with geographic information system or other mapping data to facilitate estimation and verification of travel costs where patients are unable to estimate distances (Siedner et al., 2013), and retrospective records review can also combat recall bias in the case of frequent health facility visits (Das et al., 2012). Information on resource use can also be matched with price data to minimize recall bias; however, in LMIC, there is much wider variation in price, and market prices may not accurately reflect the economic value of resources (Hutton & Baltussen, 2005).

There may also be a distinction in survey quality depending on the interviewer and where the interview takes place. Independent research assistants may be preferable to nurses if the subject material is sensitive. Individual income and spending can be sensitive, and patients may be inclined to under-report or over-report income if the purpose of the interview is not well understood (Morris et al., 2000). Using trained interviewers who understand the principles and rationale for collecting patient costs also substantially affects the quality of the data; for example, the MERGE study initially experienced poor data quality, which improved after retraining.
interviewers. Similarly, the location of the interview will affect data quality; perceived privacy will impact patient recall and willingness to disclose details on income and spending.

Finally, the medium of recording will require particular consideration in LMICs. Electronic or telephone surveys may facilitate survey completion (Walther et al., 2011) but will require some further training of interviewers in data entry and security, and planning for power and connectivity issues in fieldwork.

6. DISCUSSION

Using the aforementioned four case studies, we have highlighted important considerations in measuring patient costs and income in order to estimate the impact of illness on economic vulnerability in intervention-based contexts in LMICs.

Poverty impact metrics are currently data hungry and are therefore often excluded from study surveys because of time and budgetary constraints in a research study. Going forward in these settings, economists first and foremost have a responsibility to communicate data requirements in the study design phase and advocate for the collection of patient cost data as an essential part of the economic evaluation. Additional information on patient costs and the poverty impact of health spending is more costly to collect, but these forms of analysis are increasingly important to policy makers and program planners and therefore have a high value of information.

Inevitably, some degree of variation in methods will occur across studies where context and data availability vary. Economists therefore also must communicate with each other where different approaches are possible or where compromise as to the gold standard of data collection may be managed. Robust reporting of data collection methods can help other researchers understand and interpret findings and facilitate standardization of methods. Our recommendations for reporting data collection methods for patient costs are summarized in Table II.

Finally, it may be possible to minimize the additional cost of collecting patient cost and poverty impact data, through restricting data needs and clarifying where alternative methods are acceptable. Several alternative methodological approaches are available, and researchers must weigh limitations of potential alternatives in their own setting. Some potential advantages and limitations of various methodological approaches are described in Figure 1. We advocate for further methodological work to investigate the means to minimize the impact of cost ingredient aggregation, cost truncation, and other forms of compromise when planning poverty impact studies in LMICs, and to investigate the external validity of results that parallel effect estimates particularly in clinical trials.

<table>
<thead>
<tr>
<th>Study planning component</th>
<th>Items for consideration</th>
</tr>
</thead>
</table>
| Comprehensiveness of survey design | • Which OOP expenditures are included?  
• What is the level of disaggregation in cost ingredients and how long is the survey?  
• Are any context-specific variables included?  
• How is income measured, and whose income is collected (i.e., personal or household income)? |
| Time frame and recall | • What is the recall period for the survey? Is it appropriate to capture all economic outcomes?  
• What is the complexity of the disease pathway? Is there resulting potential for recall bias?  
• Is there potential for cost truncation in the context of chronic disease and/or future complications?  
• Are coping strategies used to estimate the long-term economic impact of health spending?  
• What is the recall period for income measurement (i.e., current vs. pre-diagnosis)? |
| Sample size and representativeness | • What is the confidence interval and margin of error deemed acceptable?  
• If estimating impoverishing expenditures, what is the distribution of pre-diagnosis income relative to the poverty line?  
• Are any adjustments to sample size required to account for clustering or non-response? |
| Data sources and survey administration | • Is a cost diary or recall used to capture expenditures?  
• Are data supplemented with any additional data sources, such as retrospective records review or GIS data?  
• Where is the interview conducted and by whom?  
• What is the medium of collecting and recording data (i.e., electronic, paper, or telephone surveys)? |

GIS, geographic information system.
Figure 1. Potential advantages and limitations of alternative approaches in data collection

### Comprehensiveness of Survey Design

<table>
<thead>
<tr>
<th>Comprehensiveness of Survey Design</th>
<th>Timeframe and Recall Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What is the level of disaggregation in cost ingredients?</strong></td>
<td><strong>What is the complexity of the disease pathway?</strong> Does the survey attempt to capture all complexity?</td>
</tr>
</tbody>
</table>
| More Disaggregated:  
  Longer survey; higher cost of implementation; potential improved response on certain variables; possible survey fatigue | More complex disease pathway:  
  Longer survey; possible survey fatigue and likely more expensive to carry out |
| Less Disaggregated:  
  Shorter survey; lower cost of implementation; aggregate cost estimates potentially less accurate less potential for survey fatigue | Less complex disease pathway:  
  Shorter survey; less potential for survey fatigue and likely less expensive to carry out |

<table>
<thead>
<tr>
<th>Are any context-specific variables included (i.e., adaptation of cost ingredients to reflect local circumstances)?</th>
<th>What is the length of the recall period for cost data?</th>
</tr>
</thead>
</table>
| Context-specific variables included:  
  Includes all important variables, survey may be less transferrable to other settings | Longer recall period: More likely to capture all relevant costs; possible recall bias |
| Context-specific variables not included:  
  May exclude important variables; may underestimate true economic burden of the illness | Shorter recall period: Possible exclusion of relevant costs; less likelihood of recall bias |

| Are coping strategies measured? | Repeated interviews with shorter recall period:  
  Less potential for recall bias, more costly to implement, may be difficult to track respondents |
|----------------------------------|----------------------------------|
| Coping strategies measured:  
  Longer survey; may more accurately capture long-term economic impact | Coping strategies not measured:  
  Shorter survey; potentially missing long-term economic impact |

<table>
<thead>
<tr>
<th>Whose income is collected (i.e., personal or household income)?</th>
<th>How does the survey attempt to capture future costs in the context of chronic disease and/or complications?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal income:</strong> Possible respondent bias; estimation of affordability more difficult</td>
<td>Assume linear trend: Potentially missing long-term economic impact, especially indirect costs and changes in income</td>
</tr>
<tr>
<td><strong>Household income:</strong> Possible breach in confidentiality</td>
<td>Frequent follow-up survey: Survey more expensive to carry out; less risk of cost truncation</td>
</tr>
<tr>
<td>GDP per capita: Poor understanding of household income relative to poverty line</td>
<td>Follow-up survey much later: Difficult to track original respondents; possible recall bias</td>
</tr>
<tr>
<td>Asset score with mapping to an absolute wealth metric: May not be reflective if population of interest is different from national population</td>
<td></td>
</tr>
</tbody>
</table>

### Sample Size

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Data Sources &amp; Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What is the confidence interval and margin of error deemed acceptable for estimation of main outcome?</strong></td>
<td>Is a cost diary or recall used to capture expenditures?</td>
</tr>
</tbody>
</table>
| Narrower confidence interval:  
  Larger sample necessary; survey more expensive to carry out. More feasible if survey is short. | Recall:  
  More feasible in context with high likelihood of literacy; higher potential for recall bias |
| Wider confidence interval:  
  Smaller sample necessary; survey less expensive to carry out. Where disease pathway is complex, may be more practical to conduct a longer survey with fewer people. | Cost diary:  
  Less potential for recall bias; more demanding of patients and requires literacy; may be costly for researchers to facilitate / encourage completion |
| How does patient follow-up affect sampling requirements? | Is data supplemented with any additional data sources? |
| Cross-sectional:  
  Larger sample is required as variation within an individual is not controlled | GIS mapping:  
  Potentially facilitate estimation and verification of travel costs; assume that traffic and road conditions remain constant |
| Cohort:  
  Smaller sample is necessary as variation within an individual is controlled for; higher risk of loss to follow-up | Retrospective records review:  
  Potentially facilitate estimation and verification of health facility visits; data may be poor or non-existent in LMIC |
| Are any adjustments to sample size required to account for clustering or non-response? | Resource use matched with price data:  
  Minimises recall bias and facilitates shorter surveys; price data may be difficult to find and/or widely varying |
| Clustering:  
  Likely especially in trial settings where respondents are randomized by health facility | |
| Non-response:  
  Those with a stigmatizing condition may be less likely to respond; response likelihood may be linked to income | |
| If estimating impoverishing expenditures, what is the distribution of pre-diagnosis income below the poverty line? | Where is the interview conducted, and by whom? |
| No patients below poverty line: Impoverishing expenditures easy to measure | **In the home vs. health facility:**  
  Weigh likelihood of potential non-response or inaccurate answers due to issues of confidentiality and/or stigma |
| Some patients below poverty line:  
  Impoverishing expenditures may inaccurately reflect full economic impact for those below poverty line; other indicators of economic impact recommended | Independent interviewers vs. nurses:  
  Weigh likelihood of potential non-response or inaccurate answers due to understanding of survey purpose, confidentiality, and cost of survey |
| All patients below the poverty line: Other indicators of economic impact are necessary | |
| **Electronic surveys (i.e., tablets):**  
  May facilitate survey completion and data quality; requires further training of interviewers and planning for power / connectivity issues | **Paper surveys:**  
  More likely to be lost; may be easier to fill for interviewers with less education |
| **Telephone surveys:**  
  Easier to track patients; may reduce issues of confidentiality and/or stigma | **Telephone surveys:**  
  Easier to track patients; may reduce issues of confidentiality and/or stigma |

© 2016 The Authors. Health Economics published by John Wiley & Sons Ltd.  
Health Econ. 25(Suppl. 1): 42–52 (2016)  
DOI: 10.1002hec
This supplement confirms the increasing implementation and sophistication of economic evaluation in LMICs (Vassall et al., 2016; Pitt et al., 2016). Going forward in these settings, evaluations need to tackle policy concerns around equity and poverty. Researchers should be challenged to address fundamental data gaps for measuring the impact of illness on economic vulnerability through stronger reporting of methods and further methodological work.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORIGINAL PUBLICATION

This work draws on experience from case studies, which have been previously published; however, all methodological recommendations and other work associated with this letter are original and have not previously been published.

ETHICS STATEMENT

No ethical approval was required for this work, as it draws only on secondary data. All case studies presented in this letter obtained ethical approval for data collection; ethics statements for these studies can be found in their corresponding research articles. All case studies gave approval for their findings to be included in this letter.

ACKNOWLEDGEMENTS

The authors are grateful to the study teams for the MERGE, XTEND, REMSTART, and ECONPOP trials for use of their data and for their insight on challenges encountered in data collection. The MERGE study team includes Tendesayi Kufa, Piotr Hippner, Salome Charalambous, Katherine Fielding, Alison Grant, and Gavin Churchyard. The ECONPOP study team includes Johanne Sundby, Katerini T. Storeng, Ouattara Fatoumata, Hanne Lichtwark, Drabo Seydou, Ramatou Ouédraogo, Patrick Ilboudo, and Torsvik Gaute. The XTEND study team includes Susan Cleary, Lucy Cunnama, Gavin Churchyard, and Edina Sinanovic. The REMSTART study team includes Sode Matiku, Bernard Ngowi, Duncan Chanda, Sokoine Lesikari, Christian Bottomley, Saidi Egwaga, Amos Kahwa, Peter Mwaba, Sayoki Mfinanga, and Shabbar Jaffar. In addition, the authors are grateful to Catherine Pitt and Ulla Griffiths for their comments on an early draft of this letter.

REFERENCES

64th World Health Assembly. Sustainable health financing structures and universal coverage. 2011: 4–7.


Foster N, Vassall A, Cleary S et al. 2015. The economic burden of TB diagnosis and treatment in South Africa. *Social Science and Medicine* 130: 42–50. DOI:10.1016/j.socscimed.2015.01.046


