Lessons Learned Document

Thematic Area	Impact and outcome evaluations
Description	This Lessons Learned document reviews recent experience documenting and attributing changes in under-five mortality to the specific interventions of a variety of iCCM programmes
Organizations documenting Lessons Learned	Malaria Consortium, Johns Hopkins University

Background

iCCM programmes intervene to directly prevent deaths from the most common life-threatening diseases in resource-constrained communities. As such, it might be assumed that iCCM programmes will result in lower mortality rates. However, in order for this to be demonstrated, three sets of assumptions must be met. First, the programme must be delivered at an **intensity** sufficient to generate impact at a population level. iCCM programme theory of change indicates that in order to generate mortality impact, there must be a substantial change in the proportion of sick children in the target population receiving safe, effective and timely treatment; in order for treatment coverage to increase, utilization of the new providers must be high and their service quality reliably adequate. Furthermore, the number of iCCM providers deployed must be sufficient to have substantially increased the overall density of service delivery points. These basic preconditions have not always been met. Second, iCCM programme design must be **appropriate for the context,** including treatment for all of the most important life-threatening conditions, medicines that are locally effective, effective targeting to those children who are truly at risk of dying and a relative scarcity of alternative providers. This set of assumptions has also not universally been met, with many programmes continuing to use co-trimoxazole for the treatment of pneumonia, for example, in spite of ample evidence of bacterial resistance to co-trimoxazole. Third, the **methods of assessing mortality** impact must be reliable, precise and generalizable. This third assumption in particular is assessed in this paper.

Process for documentation

A systematic review of published literature on the mortality impact of iCCM programmes identified one previous review (Christopher et al., 2011), summarizing information from one randomized controlled trial and several very small-scale pilots. We contacted the implementers of all large-scale programmes and research studies undertaken in Africa since 2000. For the purposes of this specific review, we limited our scope to studies with mortality data both from the period of programme implementation itself *and* from a pre-programme period of the same duration and the same seasonality. In addition, we only considered studies for which there was a comparison (non-programme) group with documented mortality experience over exactly the same time periods (in order to draw inferences about the net impact of iCCM programmes over and above secular trends in mortality resulting from other interventions). We did not consider studies in which only a handful of providers had been trained, regarding them as unlikely to be informative about the impact of large-scale national or sub-national programmes.

We identified four large-scale research studies and seven programme evaluations, from a total of eight countries, for analysis. The investigators of these studies provided standardized evaluation data for re-analysis. We defined the beginning of the implementation period as the time when at least 80% of iCCM providers had been trained and were actively offering services. The end of the implementation came with the suspension of the programme or the completion of the mortality data collection, whichever occurred earlier. The outcome considered in all cases was the mortality of children aged from 2 months to 5 years of age, measured as deaths per thousand person-years at risk. Outcomes were summarized using standard meta-analysis and meta-regression techniques, with sensitivity analyses using a variety of different weighting schemes.

Strategies that worked well

The following strategies were identified as those that helped attribute changes in mortality to iCCM programmes.

- The collection of mortality data using **full birth histories** is a promising approach for the evaluation of iCCM programmes. This is because birth history data, collected at a single moment in time towards the end of the programme implementation period, permit the reconstruction of the evolution of mortality in the target population over at least the previous two decades, with the possibility of zooming in on specific periods. The ability to understand what was happening to mortality before the introduction of the iCCM programme aids the interpretation of the study or evaluation findings, and the fact that both pre-implementation and implementation data are collected from the same households has desirable statistical properties.
- In the research studies, the intervention was introduced in a **mosaic of small geographic areas** rather than a few large areas such as districts; these intervention 'clusters' were then compared to a similar number of non-programme or 'comparison' areas. This strategy permitted relatively straightforward inference about the likely impact of the same intervention across a larger population. It also generally resulted in intervention and comparison groups starting at similar levels of pre-programme mortality.
- Some case studies included the collection of a **comprehensive dataset** including not only mortality but also changes in treatment coverage (for both iCCM and non-iCCM interventions) and detailed programme utilization data. This greatly facilitated the interpretation of the mortality findings.

Strategies that did not work well

The following strategies were identified as those that hindered the ability to attribute changes in mortality to iCCM programmes.

- Collecting birth histories in minimally literate populations requires **intensive supervision** of the data collection process, which was not achieved in all cases. Since it is difficult to detect poor quality mortality data after it has been collected, we relied entirely on an assessment of the plausibility of the levels of mortality assessed at baseline; it is likely that some moderately poor quality data passed this insensitive test. Indirect methods of mortality assessment are not useful because they cannot be mapped to a precise historical period, and nor can a precise age range be specified. Prospective demographic surveillance is not always well positioned to capture deaths of the youngest infants.
- There were **basic flaws in the evaluation design** of the vast majority of studies, making it very difficult to draw any inference from the results. Either the comparison areas were vastly different from the interventions at baseline, and/or the programme was allocated to very small numbers of relatively large geographic areas, making it impossible to rule out the influence of idiosyncratic local changes on the findings, and/or the comparison areas also received some form of iCCM or

similar intervention. In several cases, the programme delivered in the intervention areas was so different from the standard model of iCCM that the value of comparing across studies has to be questioned.

- Because mortality is a rare event, virtually none of the cases studies was actually **powered to be able to detect a statistically significant impact** of the programme. Power calculations, which are a basic step in the development of an evaluation plan, were either simply not done, were based on out-of-date or wildly over-optimistic assumptions, or did not take the evaluation design into account. In addition, programme exposure periods were either far too short to accumulate sufficient numbers of deaths in the study population, and/or did not give the targeted populations time to get used to using the new providers. Thus, *huge amounts of evaluation resources were wasted on studies which never had any possibility of achieving their principal aim of demonstrating programme impacts on mortality.*
- The programmes took place in areas with **rapidly evolving health systems and epidemiological contexts**. Thus, they often no longer met the basic assumptions required to demonstrate mortality impact.

Lessons Learned

- **Mortality impacts in recently implemented programmes vary considerably**, from a (statistically significant) 76% reduction in mortality, to a (non-significant) 43% increase, with a median reduction of 21%. We believe that much of this apparent variation is due to inappropriate programme and evaluation design.
- Mortality measurement is a very specialized activity requiring well-trained interviewers, close supervision of fieldwork, and—above all—enormous survey sample sizes. **Funders and government sponsors of iCCM evaluations should not require direct mortality assessment unless they are prepared to invest the requisite resources.**
- There is a logical chain of iCCM results in which high utilization of quality services is the precondition for high coverage of safe, effective, and timely treatment of sick children, and high coverage of safe, effective, and timely treatment is in turn the precondition for observing reduced mortality. Mortality measurement should not be undertaken unless it can be demonstrated that the other preconditions have already been met. In general, **mortality outcomes should probably not be considered for any programmes likely to implement for less than two years.**
- Because of rapidly changing health systems and epidemiological contexts**, it is much easier to interpret mortality data if companion data on treatment coverage, programme utilization, and other contextual variables are also collected**.
- **Programme design considerations often conflict with the basic premises of good evaluation.** For example, one programme decided to introduce iCCM in two districts and compare their mortality experience with one comparison district. This "2 versus 1" comparison is known in the evaluation literature to produce results which are uninterpretable. Likewise, pushing programmes into the highest mortality districts inevitably means that comparison areas will not be truly comparable at baseline, creating extreme difficulties in interpreting evaluation findings later on.
- If a strong evaluation context can be guaranteed, **full birth histories—or, better still, full *pregnancy* histories—are the ideal way of collecting data on child mortality**. They should be analyzed by compartmentalizing both deaths and person-years at risk into multiple sequential time periods. A single birth history survey can produce information both for the programme implementation period and for the pre-programme period.