

A Comparison of Sub-national Survey Data and Health Management Information System Data in Zambia

Rose Gabert

A thesis
submitted in partial fulfillment of
the requirements for the degree of

Master of Public Health

University of Washington
2014

Committee:
Emmanuela Gakidou, PhD
Haidong Wang, PhD
Marie Ng, PhD

Program Authorized to Confer Degree:

School of Public Health
Department of Global Health

University of Washington

Abstract

A Comparison of Sub-national Survey Data and Health Management Information System Data in Zambia

Rose Gabert

Chair of Thesis Committee:

Emmanuela Gakidou, PhD

Department of Global Health

Child health outcomes and coverage of key maternal and child health interventions vary widely across districts in Zambia. Sub-national data is necessary to empower health decision makers to target informed strategies to improve service delivery and population health. Some efforts have been made to produce representative estimates of coverage and health outcomes at the district level in Zambia. This investigation explores the feasibility of using data from the Health Management Information System (HMIS) to inform the estimation of trends in child mortality (5q0), skilled birth attendance (SBA), and antenatal care (ANC). District-level trends in 5q0, SBA, and ANC were constructed using HMIS data from 2000-2008. HMIS-based trends were then compared to survey-based data using visual analysis and Pearson's correlation. Resulting correlations ranged from strongly-correlated ($r > 0.6$) to negligible association ($r = 0$) to strong negative correlation ($r < -0.6$). In some districts and indicators, HMIS data mirrored the trend produced using survey data and statistical modeling, while in others the two datasets failed to agree. Further investigation is necessary, but preliminary results suggest that HMIS data is a powerful and neglected resource for subnational estimation

©Copyright 2014

Rose Gabert

Introduction

Child health outcomes and coverage of key maternal and child health interventions vary widely across districts in Zambia and have shown dramatic change in the last decade. For example, the percentage of women aged 15-49 who received at least four antenatal care visits in 2010 pregnancies ranged from 3% in the central Kasempa district to an impressive 97% in the northern district of Mwense in that same year. Similarly, while the national proportion of women giving birth with skilled professionals (SBA) increased to 55% in 2010, the trend in some districts fell as low as 1% and climbed as high as 97% in others, with some districts showing improvement from 2000 to 2010 and others rapid decline¹. The degree of heterogeneity is such that national-level estimates and trends are insufficient to support evidence-based decision making to improve country health.

While the health gains and high levels of intervention coverage in some districts call for policies to bolster efforts and continue successful approaches, service delivery systems in other districts are clearly in need of reassessment and potential restructuring. Without detailed and timely data of subnational trends, it is impossible for policy makers and decision makers to target informed strategies that meet the needs of their populations.

The figures cited above come from a recently-published report detailing the progress of child health outcomes and key health intervention coverage across districts in Zambia. The ambitious report uses small area estimation methods and rigorous statistical modeling to synthesize over thirty nationally-representative data sources at given points in time, producing distinct district-level trends from 2000 to 2010. The resulting report on district progress and health outcomes is a powerful tool for in-country policy makers and health professionals to assess priorities in their districts and track their progress against the rest of the country.

And yet, most ministries of health and district health offices throughout the developing world lack the resources to reproduce such a computationally intensive exercise. The sheer investment of manpower required to compile, clean, and standardize over thirty large data sets, the computational capacity to handle decades of data points, and the statistical expertise to run simultaneous multiple regressions and advanced nonparametric models on specialized statistical software is out of the reach and budgets of most ministries of health and district health offices, despite their dedication and commitment to excellence.

Moreover, detailing health trends from 2000 to 2010 and published in May of 2014, the report is already markedly out of date. More timely, accessible subnational data is necessary for ministries of health, district health officers, and public health policy makers. Health Management Information Systems collect routine data on health system inputs, outputs, and outcomes at the facility level. With standardized indicators and regular systems of collection and collation, HMIS provide timely, disaggregated data on the health system and those it serves.

Researchers have traditionally struggled with several problematic aspects of HMIS data. First, varying degrees of data quality and completeness are observed across HMIS systems. A recent report by the National Bureau of Economic Research recorded 3,318 missing records out of some total 45,000 facility quarterly reports that should have been submitted to the Zambia HMIS between 2000 and 2008. Secondly, the true degree of missingness can be difficult to ascertain. When a facility fails to send in a

quarterly report all data elements for that month are missing, but if a given facility department failed to keep accurate records during the quarter, additional “invisible missingness” may persist in the data.²

Furthermore, changes in diagnostic technology, medical procedures, and regulations within the HMIS itself can problematize apparent trends in disease. An obvious example arises from the uneven scale-up of rapid diagnostic tests across districts in Zambia, which makes comparison of counts of malaria outpatient visits over time and between districts unreliable.³

Finally, researchers point out that even if record keeping and data collection were perfect and uniform, HMIS data is ultimately capturing only those members of the population that enter the facility. It is unknown if disease prevalence and coverage of health indicators would be different in the population that is absent from HMIS data. This presents a key difficulty in determining an appropriate denominator to make comparisons of raw HMIS counts across facilities and districts.^{4,5,6}

Despite these inherent and weighty challenges, HMIS data remains the only consistent source of routine national health data. If it can be shown that despite the issues described above, HMIS data can be used to track subnational trends in health outcomes and intervention coverage, the HMIS would be a powerful tool for ministries of health, district health officers, and public health professionals to make data-driven, targeted decisions to improve the health of the populations they serve.

The primary objective of this investigation is to determine whether HMIS data in Zambia can be used to track trends in key health outcomes and intervention coverage indicators at subnational levels. Specific questions include:

1. Can counts of under-5 death recorded in the Zambian HMIS be used to track survey-based trends in child mortality at the district level?
2. Can counts of deliveries recorded in the Zambian HMIS be used to track skilled birth attendance at the district level?
3. Can counts of antenatal care visits recorded in the Zambian HMIS be used to track survey-based trends in ANC coverage at the district level?

Methods

Three survey-based indicators were chosen from the report *Progress in Child Health Across Districts in Zambia* for their importance to public health, their reliability of measure, and their rate of change over the time period in question (2000-2008). The probability that a child born alive will die before age 5 years (child mortality, often denoted 5q0) is a central indicator of population health, and is more reliable than disease-specific rates of death which rely on diagnostic accuracy. The proportion of women 15 to 49 years who receive at least four antenatal care visits during pregnancy (ANC4) and the proportion of women 15 to 49 who deliver with skilled birth attendants (SBA), are considered tracer indicators of mother and child health intervention coverage.⁷ Moreover, as highlighted above and in the report itself, ANC and SBA showed marked heterogeneity across districts and over the time period in question, which make them a good metric against which to measure the ability of HMIS data to follow dynamic trends. The three survey-based indicators selected are detailed in Table 1.

Table 1. Survey-Based Estimates from Progress in Child Health Across Districts in Zambia Report¹

| Estimate | Description | Sources synthesized for estimation |
|--------------------------------|--|---|
| Child mortality (<i>s</i> q0) | Probability of dying between age 0 and 5 | Zambia National Census 1990, 2000, 2010; DHS 1992, 1996, 2007 |
| Antenatal care (ANC4) | Proportion of women 15-49 years who had four or more antenatal care visits at a health facility during pregnancy | DHS 1992, 1996, 2001-2002, 2007 |
| Skilled birth attendance (SBA) | Proportion of pregnant women aged 15-49 years who delivered with a skilled birth attendant (a doctor, nurse, midwife, or clinical officer) | DHS and MICS in 1992, 1996, 1999, 2001, 2007 |

Comparable indicators to match these three survey-based estimates were then constructed from HMIS data elements and other administrative data sets. The HMIS records facility-level under-5 deaths due to all causes, tallied from inpatient and brought-in-dead registries. In addition, service delivery aggregation forms in the HMIS aggregate services delivered by a facility in a quarter, including first time antenatal care visits, return antenatal care visits, live births, still births, normal deliveries, complicated deliveries, and Caesarean sections.⁸ A total number of facility births was calculated as a sum of live births and still births. Similarly, total facility deliveries were calculated as a sum of normal, complicated, and Caesarean section deliveries. While multiple analyses including similar data elements may seem redundant, this investigation explores the feasibility of HMIS data for this purpose, and it is yet unknown which data elements best align with survey-based estimates.

Counts extracted from the HMIS are detailed in Table 2.

Table 2. HMIS Data Elements

| HMIS Data Element | Description |
|------------------------------|---|
| Under 5 Deaths | Inpatient registries + brought-in-dead registries |
| First Antenatal Care Visits | Kept distinct for separate analysis |
| Antenatal Care Reattendances | Kept distinct for separate analysis |
| Facility Deliveries | Normal Deliveries + Complicated Deliveries + Caesarean Sections |
| Facility Births | Live Births + Still Births |

HMIS forms are submitted to the district health office by three distinct types of facilities: hospitals, health centers, and hospital-affiliated health centers (HAHCs). Thus, facility quarterly reports can be summed to produce district-platform-year data. In addition, the level of district-platform-year completeness can be calculated. Each raw HMIS count—under-5 deaths, ANC first visits, ANC return visits, births, and deliveries—was summed at the district-platform-year level and adjusted to account for

underreporting, calculated as the percent of facility-quarters that were missing in a given year. District-platform-year counts were then summed to create adjusted counts at the district-year level.

Quarterly facility-level counts of under-5 deaths, ANC visits, births, and deliveries had now been crudely adjusted for facility missingness and summed to district years. However, the problem of comparing raw count data to percentages and probabilities, the so-called denominator question, remained. 5q0 is the probability, expressed as a rate per 1,000 live births, that a child born alive will die by age five. HMIS death counts of children under 5 were divided by total births in the district and multiplied by 1000 to approximate a probability of death before age 5 given birth. District-level births in 2000 and 2010 were extracted from the Zambian National Census and interpolated for intervening years.

Similarly, to compare district-level counts of ANC visits, facility births, and deliveries to percentages of women receiving those services during pregnancy, the counts were divided by total births in the district, using numbers of district births interpolated from the Zambian National Census.

The final five pairs of indicators to be analyzed are detailed in Table 3 below.

Table 3. Survey-Based and Administrative Indicator Pairs

| Survey-Based Estimate | Administrative Data Comparator |
|--------------------------------|--|
| Child mortality (5q0) | Under-5 Deaths (HMIS) / District Births (interpolated census) |
| Antenatal Care Coverage (ANC4) | First Antenatal Attendances (HMIS)/District Births (interpolated census) |
| | Antenatal Reattendances (HMIS)/District Births (interpolated census) |
| Skilled Birth Attendance (SBA) | Facility births (HMIS)/District Births (interpolated census) |
| | Facility deliveries (HMIS)/District Births (interpolated census) |

Finally, having constructed simple yet comparable HMIS indicators of 5q0, ANC, and SBA coverage, a method of comparison was required. As this investigation is interested in the suitability of HMIS data to track trends rather than provide precise estimates, directionally-specific correlation between a pair of indicators seemed the logical metric of comparability. Within each district and indicator pair, yearly survey-based estimates were regressed on the HMIS-constructed complement using standard ordinary-least squares regression and Pearson correlations were compared between districts and indicator pairs. All analyses were performed using Stata 13.⁹

The distribution of correlations within each indicator pair across all 72 districts were visualized using histogram figures. These five distributions were used to determine a ranking system of 1-5 for each district indicator, outlined in table 4 below.

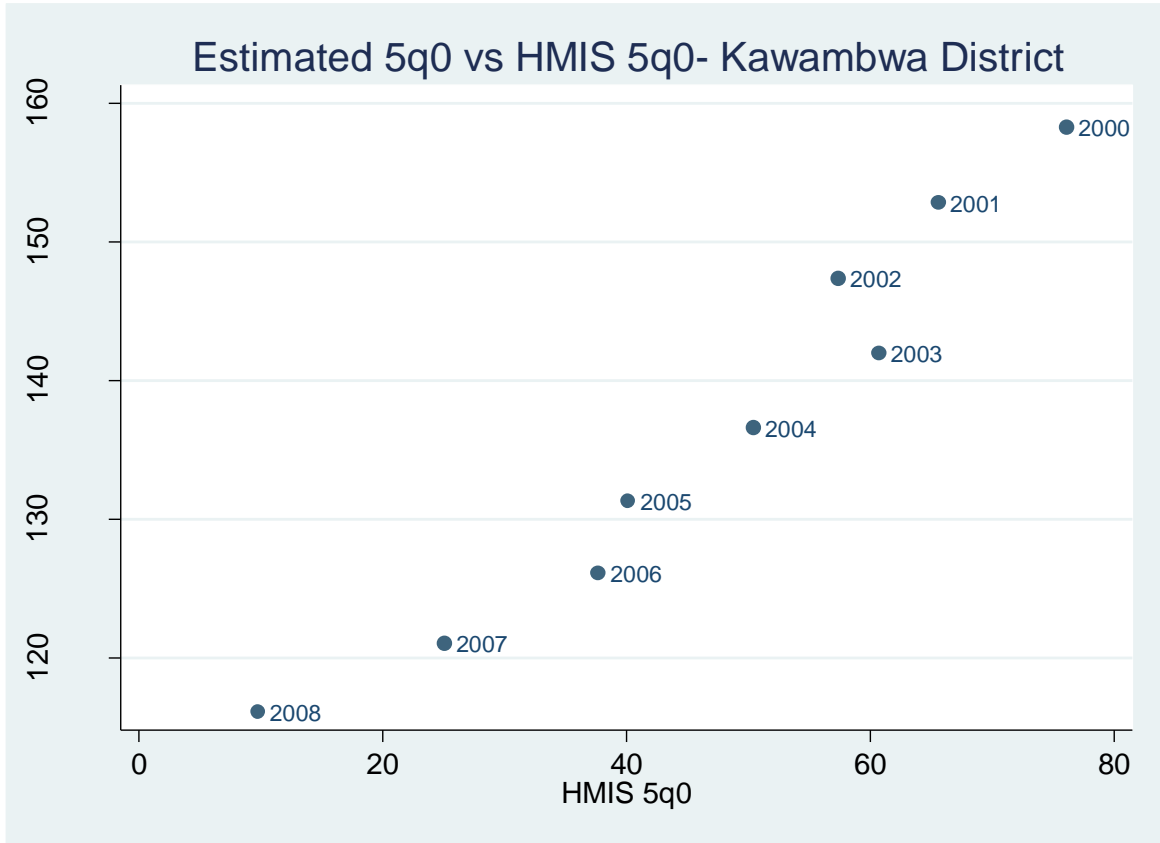
Table 4. Rankings of District-Indicator Pair Correlation

| Rank | Correlation Range |
|------|-------------------|
| 1 | 0.6 < r ≤ 1 |
| 2 | 0.3 < r ≤ 0.6 |
| 3 | -0.3 < r ≤ 0.3 |
| 4 | -0.6 < r ≤ -0.3 |
| 5 | -1 < r ≤ -0.6 |

Results

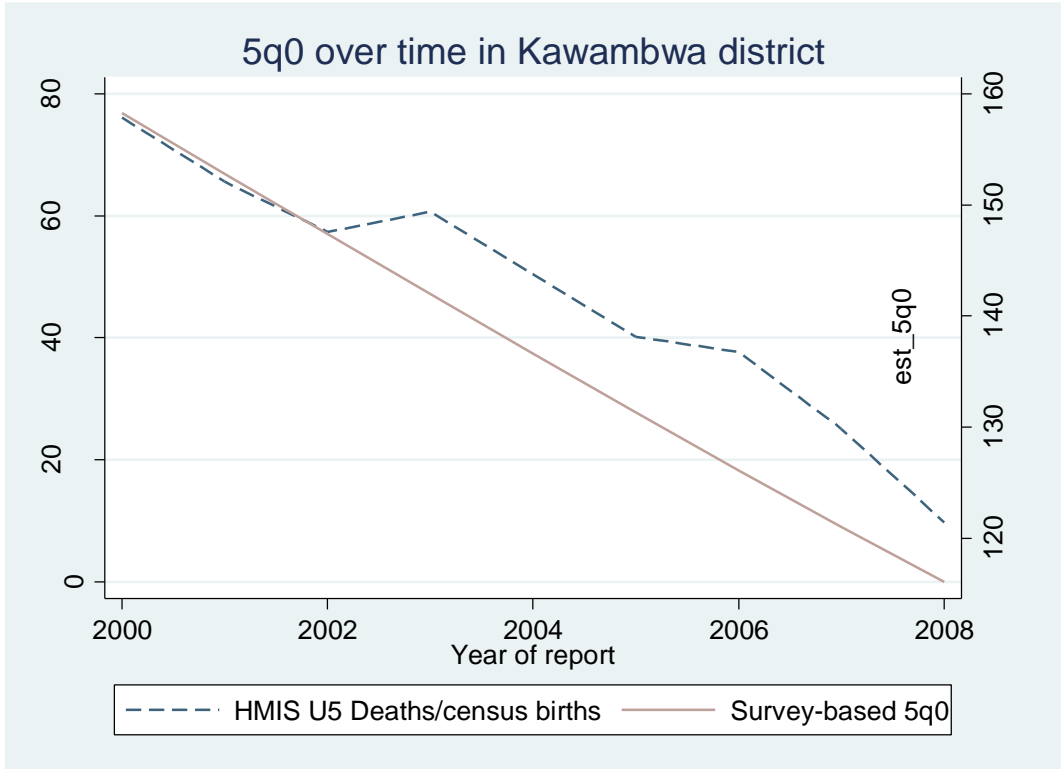
Correlation between survey-based estimates and HMIS data was dramatically varied across districts and indicator pairs. While some districts and indicators showed remarkable comparability of trends between survey-based estimates and HMIS data elements, the so-called Group 1 indicator-districts, the data revealed no relationship or negative correlations between other district-indicators. Figure 1 below exemplifies the pattern seen in 5q0 comparisons in Group 1. A regression of survey-based and HMIS estimates of 5q0 resulted in an r value of 0.954 in Kawambwa district.

Figure 1. Yearly survey-based 5q0 estimates vs HMIS deaths/census births in Kawamba district



While the Y and X axes illustrate the degree of difference between the point estimates (an order of magnitude) the trend is steady and nearly-perfectly correlated. As such, a graph of both trends over year, as shown in figure 2, shows agreement between the two estimation methods as child deaths fall from dramatically between 2000 and 2008.

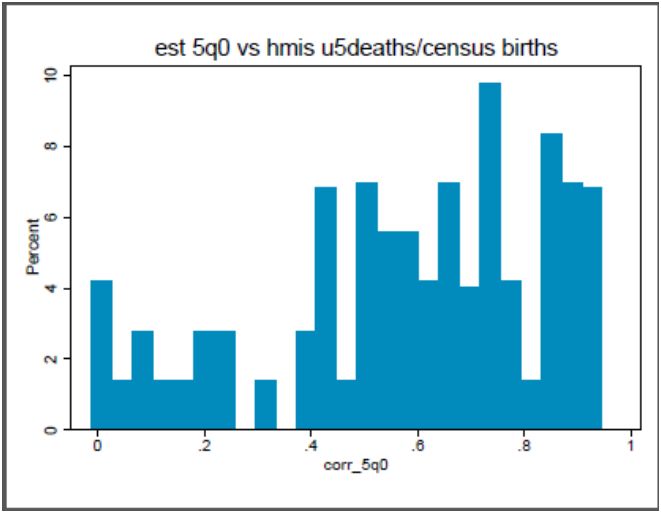
Figure 2. Survey-based and HMIS child mortality over time



The HMIS-derived estimates of 5q0 exhibit more noise than the smooth trend predicted by survey-based estimates, and yet the trends are markedly similar.

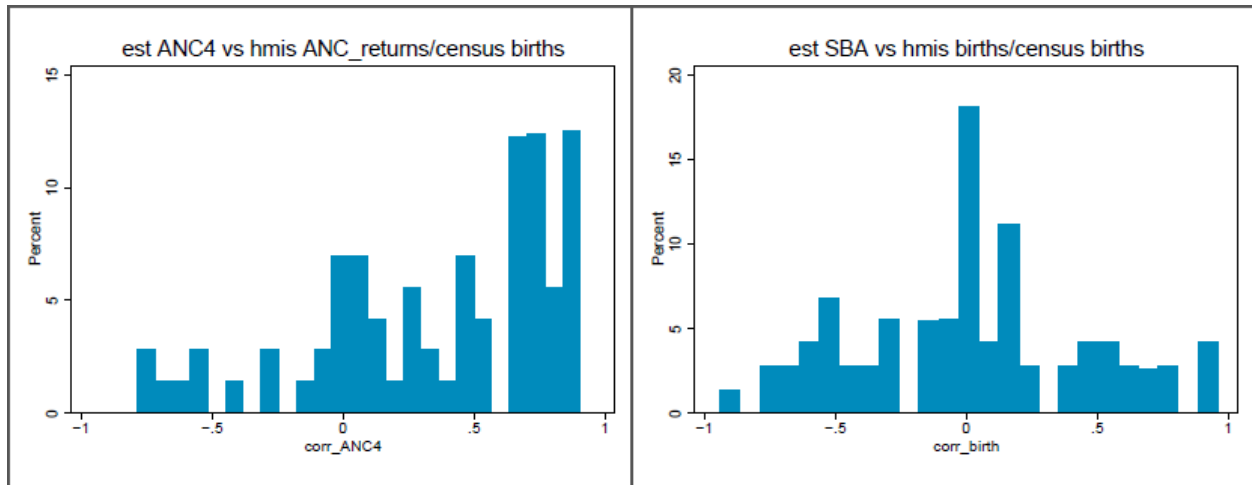
Across districts, correlations between survey-based estimates and HMIS-derived estimates were narrowest in the deaths and child mortality pairing, within which only 3 districts were negatively correlated. The distribution of correlations is shown in Figure 3 below.

Figure 3. Distribution of district correlations between survey-based 5q0 estimates and HMIS U5 deaths/census births



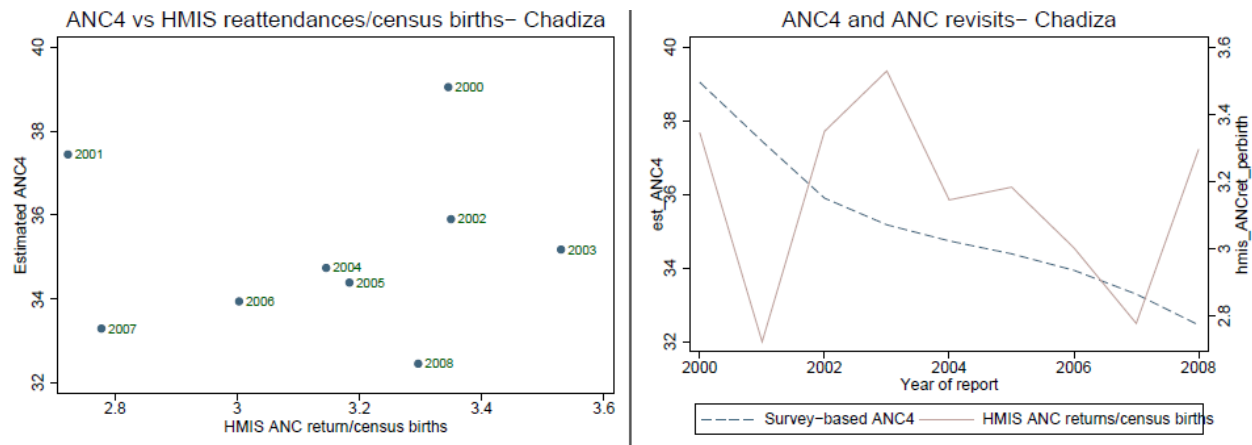
In contrast, the ANC, births, and deliveries indicator pairings varied greatly across districts, with a marked number of districts exhibiting no pattern or negative correlations below 0. Figure 4 below illustrates the distribution of district correlations between survey-based ANC4 coverage estimates and HMIS ANC return visits divided by district births as well as the distribution of correlations between SBA and HMIS facility births divided by district census births.

Figure 4. ANC and SBA district correlation distributions.



In comparison to the strong correlation seen in Group 1 Districts such as Kawambwa, above, many districts showed limited or negative correlation across the ANC and SBA indicator pairings. Figure 5 below is an example of the lack of pattern seen in Group 3 ANC4 Districts.

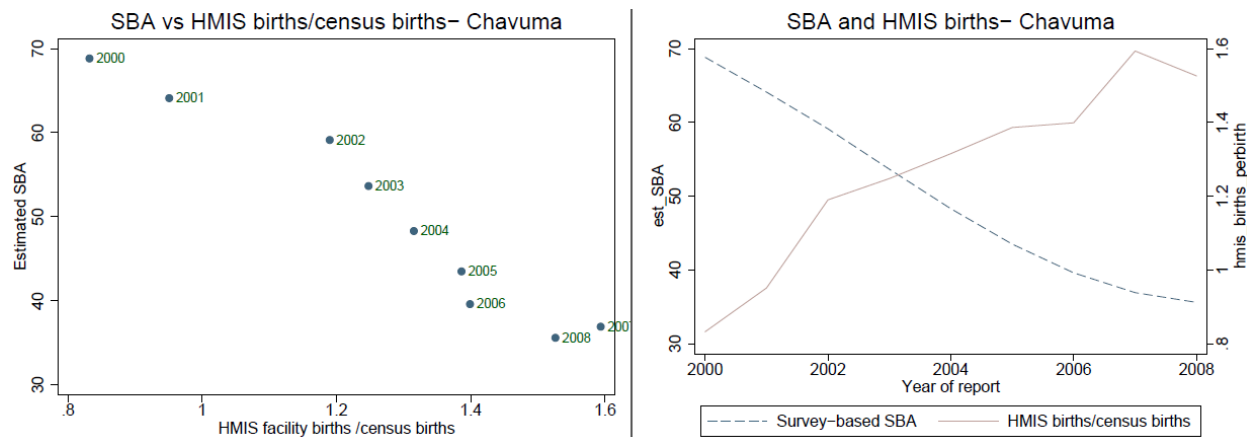
Figure 5. Survey-based ANC4 estimates vs HMIS ANC revisits/census births in Chadiza District



While some data points on the left-hand graph of Figure 5 seem to indicate a positive relationship between the two trends, other data points stray. The correlation of ANC4 and HMIS-derived antenatal coverage is 0.007 in this district. On the right-hand side the survey-based ANC trend shows a smooth decline in ANC coverage over time while the solid HMIS ANC returns visits indicator swings wildly between 2000 and 2008.

While some indicators and districts had strong relationships (Group 1) and others had indistinct associations or relatively no pattern (Group 3) still others showed strong negative correlations. Figure 6 below uses births in Chavuma district to illustrate the negative correlation patterns of Group 5.

Figure 6. Survey-based SBA vs HMIS births divided by district births in Chavuma district



In stark contrast to Kawambwa District under-5 deaths shown in figures 1 and 2 above, the correlation between survey-based SBA and HMIS births divided by district births in Chavuma district is almost perfectly negative, with r equal to -0.941 . On the right-hand panel of figure 6 the survey-based estimate shows a smooth decline in SBA coverage from 2000 to 2010 while the HMIS data shows increased return ANC visits per district births over the same time period.

The extreme heterogeneity of correlations across districts and indicator pairs makes a summary statistic or graphic difficult. A heat map of the 72 districts and five indicator pairings explored, the attached Figure 7, gives a sense of the distribution and variation of correlations.

Discussion

This investigation found striking variation in the comparability of survey-based estimates and HMIS-derived indicators of child mortality, antenatal care coverage and skilled birth attendance. The data did not suggest that HMIS numbers underestimate or inflate survey-based estimates across districts in any systematic way. A universal scaling or inflation factor could not be determined by indicator pair or by district. No districts proved to be categorically well- or poorly-correlated across all indicators.

And yet, despite the missingness, denominator, bias, and consistency concerns often cited as evidence against using HMIS data, a number of districts and indicator pairs performed markedly well in this analysis. Over 50% of districts demonstrated correlations over 0.6 comparing crude counts of under 5 deaths in HMIS facilities divided by interpolated census births to the most recent survey-based estimates.

Indeed, the heterogeneity of correlation across districts and indicators suggests a more thorough consideration of the assumption upon which this paper is predicated: that survey-based estimates of health outcomes and coverage are a gold-standard against which to measure HMIS data. The following paragraphs explore the strengths and limitations of administrative data and survey-based measures for health outcomes, child mortality, and coverage estimates, SBA and ANC4. In particular, the discussion

explores the possibility that HMIS and other administrative data be a complement to intrinsic weaknesses of survey-based estimates.

Health outcomes—child mortality

In the evaluation of survey-based and administrative estimates of health outcomes it is important to be precise about the performance of a given aspect of estimation; the question of completeness, the proportion of events that are being captured in a given point-estimate, is separate from questions of point-estimate or time trend accuracy. Moreover, child mortality and intervention coverage are generated metrics composed of a numerator and a denominator which may have differential biases and limitations. The strengths and weaknesses of survey-based birth, death, and child mortality estimates have been described, examined, and refined elsewhere.^{10 11}

Survey-based estimates of births, deaths, and child mortality are certainly more complete point-estimates than the administrative data. A glance at the differential scales in Figure 2 shows that despite strikingly parallel trends, survey-based numbers estimate that 2-4 times as many children die in Kawambwa district than are captured by the HMIS data. This is largely due to the fact that the HMIS is only counting those deaths that occur in the facility or are reported there. We assume that many if not most children die outside of facilities.¹

This is in fact an important limitation of the facility data, the implications of which are manifold. Primarily, the HMIS data is simply blind to a percentage of numerator events and is therefore underestimating the probability of those events. In addition, as a population estimate, HMIS-based child mortality is capturing only the experience of those individuals that enter the facility, a population that may be more or less affluent and healthy than the population at large.

A similar yet contrasting bias occurs in survey-based estimates of child mortality, which rely on the survival of a woman to be interviewed regarding her fertility and child-death experience. Assuming that mothers' and children's survival is correlated, which is highly likely in a population characterized by high infectious disease burden, the estimate is strongly biased downward.¹⁰ HMIS data, though victim to its own ascertainment bias of those who enter the health system, records events that happen in that facility regardless of mother's survival.

The temporal certainty of HMIS data is perhaps its greatest strength, even when point-estimates, rather than the trends over time, are of primary concern. The summary and complete birth history methods most commonly used in survey-based estimation of 5q0 utilize some iteration of the Brass Method, which describes a fertility and mortality experience of mothers and children that is not precisely calibrated to the time of the survey but is an indirect estimate of the combined experience of cohorts of women based on past conditions.¹¹ The survival bias inherent in survey-based child mortality estimates may change rapidly over time with the natural course of epidemics, particularly in populations with high mortality due to HIV/AIDS. In addition, fertility biases created by the Brass Method spread the estimated proportion of children dead across a cohort of women and over an estimated time period. In a population with high mortality due to infectious disease, and in a time period such as 2000-2008, during

¹ It should be noted that it is likely that the interpolated census births used here as a denominator underestimate total births. It is likely that the HMIS child mortality estimates are even less complete and this estimated 5q0 is spuriously inflated by an incomplete denominator.

which ART, PMTCT, ITNs, IRS, IPTp, and ACTs were rapidly scaled up [CITATION], this is a not insignificant source of bias. In contrast, HMIS-recorded deaths are precisely pinned to the year in which they occurred, are sensitive to rapid changes in mortality characteristic of epidemics, and should reflect the changing impact of the health system as it responds to meet population needs.

The demographic methods used in survey-based child mortality estimation are inherently imprecise over time and include a survival bias that may be significant in populations of high infectious disease mortality; this analysis suggests that standard estimates of child mortality from one survey could be improved by utilizing a model-based approach that includes some HMIS-based metric of child mortality. The HMIS data could help synchronize the time period of interest and limit bias introduced through indirect estimation, yielding more accurate point-estimates in each given year.

Furthermore, the survey-based estimates used in this investigation collapsed multiple surveys over multiple years and borrowed information across time and geographic location. Including HMIS data could help improve inaccuracy introduced not only by indirect demographic methods, but also due to over-smoothing between adjacent time periods or districts. The HMIS numbers are pinned to the geographic location and year, including them could make the model more sensitive to realities on the ground.

However, the problem of missingness in HMIS data cannot be ignored. It is possible that more deaths are counted as time goes on and reporting improves. In addition, individual hospitals or health facilities may drop out of reporting or undergo other data quality issues, leading to spurious noise. Further investigation into HMIS missingness, particularly the question of known versus unknown missingness, and its implications for the usability of HMIS data after multiple imputation, is required. Regardless, this investigation raises interesting questions regarding the possibility of using HMIS or other administrative data to augment the accuracy of demographic methods currently in wide-use for estimating child mortality from survey data.

Coverage estimates—SBA and ANC4

The feasibility of using HMIS data to improve estimation of coverage indicators has parallels to and interesting departures from child mortality estimation. Similar to the above discussion, biases and limitations may have differential implications for numerators and denominators, as well as for completeness, point accuracy, and trend accuracy in these data sets.

Theoretically, correlation between HMIS data should be better with SBA and ANC4 coverage than with child mortality because the numerator event is happening almost exclusively in facilities. The inherent bias of HMIS data, that it counts only those individuals entering the health system, does not apply. Moreover, the question of completeness is largely irrelevant, because unlike the HMIS underestimating child deaths, nearly all ANC visits and births with a skilled attendant are taking place in facilities. The only underestimation issue should be in recording. And yet, the results of this investigation show that correlations are not only worse between HMIS and survey-based data in the SBA and ANC4 metrics, but in some cases show strong negative correlations. This may be due to discrepancies between HMIS-derived numbers and survey-based estimates in general or due to the modeling strategies utilized in generating these particular estimates over space and time.

Cross-sectional survey estimates of coverage are subject to their own biases regardless of modeling. There may be selection bias in populations with high mortality, as women who receive services may not survive to report them due to fundamental differences between women who are more likely to receive services and those who are not. This relationship is complicated by the correlation between receiving services and mothers' survival, which may not be insignificant in a population with high rates of HIV/AIDS and malaria. ANC visits include services that could directly impact mother's mortality due to these diseases, including preventive measures such as ITNs and IPTp, as well as screenings that could lead to life-saving treatment. Survival bias in this case would lead to over-estimation of intervention coverage.

In addition, there is a bias not just from mother's survival to time of survey but in pregnancies carried to term and resulting in live births, outcomes that are also likely to be correlated with ANC and SBA services. Women may receive these services, an ANC visit or an in-facility birth, and yet not report the event of the service because the pregnancy did not result in live birth.

In contrast to these various ascertainment biases, HMIS data records actual events. The number of births in a facility does not change if the mother does not survive to the time of survey. Similarly, the number of ANC visits per birth does not change if a proportion of pregnancies recorded by the HMIS do not result in live birth. Moreover, the concern that HMIS data is more representative of the population that has better access to facilities does not apply because the count includes all facilities and the majority of events take place in facilities; access should by definition be associated with coverage. Finally, using HMIS data directly links the estimate to variations in health system output, which should be highly correlated to intervention coverage. These considerations may make HMIS data a more accurate estimate of intervention coverage, despite the fact that the current analysis may not have settled a perfect denominator.

Finally, coverage estimates are hindered by a question of representativeness, as they are derived from questioning a subset of individuals and may not reflect the experience of the population as a whole. This issue is especially pronounced when survey-based coverage estimates are pushed to provide subnational estimates with much smaller sample sizes that are not systematically sampled. In addition, when DHS samples within a district are used to represent the district as a whole, they may be even more biased because the sample may include a few clusters, households, or women which are undoubtedly more alike to each other than to the rest of the district. HMIS data may produce estimates of SBA and ANC coverage that are not only more accurate but also more representative.

In the current investigation, survey-based estimates were modeled first with covariates and then across space and time. This method may be problematic in districts with smaller sample sizes or no data in some years. While the coverage estimates relied entirely on surveys that sampled different numbers of households from different districts over the years and omitted some districts entirely, the child mortality estimates were anchored by indirect-estimation methods in the 2000 and 2010 Zambian Censuses, which were carried out in every district with every woman. The small-numbers and lack of representation were less of an issue in the child mortality than in the intervention coverage estimates. This may explain why correlations between survey-based and HMIS indicators were consistently higher in the child mortality comparisons than in ANC4 and SBA, and the strongly negative correlations observed only in ANC4 and SBA pairings. The modeled approach found that ANC4 coverage went down on average, and this average may be reflected in those districts that had little data.

While more sophisticated demographic and statistical methods could be used to construct appropriate denominators for HMIS data, such as using DHS data to estimate age-specific fertility rates and thus total numbers of births, the original spirit of this investigation was to compare purely administrative data sets available with limited or no statistical modeling to more rigorous, survey-based estimates produced with statistical methods. As such, dividing under-5 deaths, counts of ANC visits, and counts of facility births and deliveries by crudely-interpolated census birth counts balanced the desire for simplicity with a need for comparability to the survey-based indicators of 5q0, ANC, and SBA.

However, if we relax the assumption that survey-based estimates are a gold standard against which to compare other data sources and methods, this analysis raises interesting implications for the usefulness of HMIS data as an analytical tool in its own merit. It is possible that a more refined HMIS-based indicator would prove useful in augmenting survey-based estimates, a supposition that merits further investigation.

Sticking to the original assumption, that survey estimates are the most precise and reliable method of estimating child mortality and intervention coverage, the fact remains that the ambitious undertakings of survey-based subnational estimation such as the recently-published report Progress in Child Health Across Districts in Zambia remain out of the reach of ministries of health, district health offices, and planners and policymakers the world over. Large national population-based health surveys do not yield information on key health outcomes and intervention coverage quickly enough nor at sufficiently disaggregated levels that can be used to assess and compare trends across geographies within a country. This analysis demonstrates the as-yet-unrealized possibility of HMIS data to fill this urgent gap.

Limitations

A large limitation of this study is the inclusion of only one country. Sub-national survey-based estimates of child mortality and intervention coverage are currently limited, but as they become available comparison to HMIS numbers will be interesting.

A similar investigation of multiple countries aggregated to national levels might prove enlightening not only in expanding generalizability but also in isolating biases and limitations of HMIS and survey-based methods. The added complications of collapsing multiple estimates, small numbers, borrowing across space and time, and representativeness dilute what may be a stronger relationship between the survey-based and HMIS estimates on their own.

In addition, the implications of this study for estimating mortality are limited by the narrow focus on the under-five population. Reporting and facility capture are likely very different in older populations.

This study used data from the Zambia HMIS up to 2008 when the system switched to new and likely improved procedures and data collecting methods. An updated analysis would generate more relevant and possibly more robust conclusions.

Finally, the problems posed by missing data in the HMIS cannot be understated. In light of this analysis, HMIS data could be a powerful asset in survey-based health outcome and intervention coverage estimation, especially at subnational levels. Further investigation into the reliability of HMIS data across different data elements, the relationship of HMIS missingness to other non-HMIS health indicators, and the feasibility of various imputation methods is required.

-
- ¹ *Assessing Impact, Improving Health: Progress in Child Health Across Districts in Zambia*. Seattle: Institute for Health Metrics and Evaluation; 2014. <http://www.healthdata.org/policy-report/assessing-impact-improving-health-progress-child-health-across-districts-zambia>.
- ² Ashraf, Nava, Gunther Fink David N Weil. Evaluating the Effects of large scale health interventions in developing countries: the Zambian Malaria Initiative. Working paper 16069. Cambridge, MA: National Bureau of Economic Research, 2010.
- ³ Chizema-Kawesha Elizabeth, John M Miller, Richard W Steketee, Victor M Mukonka, Chilandu Mukuka, Abdirahman D Mohamed, Simon K. Miti, Carlos C Campbell. Scaling up malaria control in Zambia: Progress and Impact 2005-2008. *Am. J. Trop. Med. Hyg.*, 83(3), 2010, pp. 480–488
- ⁴ McGinn Therese, Judy Austin, Gace Kodindo, Tim Ensor. Trends in Maternal Outcomes and Utilisation in Zambia: evidence from secondary sources. Lusaka, Zambia, 2010.
- ⁵ Etyang O Anthony, Kenneth Munge, Erick W Bunyasi, Lena Matata, Carolyne Ndila, Sailoki Kapesa, Maureen Owiti, Iqba IKhandwalla, Andrew J Brent, Benjamin Tsofa, Pamela Kabibu, Susan Morpeth, Evasius Bauni, mark Otiende, John Ojal, Philip Ayieko, Maria D Knoll, Liam Smeeth, Thomas N Williams, Ulla K Griffiths, J Anthony G Scott. Burden of disease in adults admitted to hospital in a rural region of coastal Kenya: an analysis of data from linked clinical and demographic surveillance systems. *Lancet Global Health* 2014;23216-24.
- ⁶ Moisi JC, Nokes DJ, Gatakaa H, et al. Sensitivity of hospital-based surveillance for severe disease: a geographic information system analysis of access to care in Kilifi district, Kenya. *Bull World Health Organ* 2011; 89: 102–11.
- ⁷ Monitoring maternal, newborn and child health: understanding key progress indicators. Geneva: Switzerland; World Health Organization(WHO), 2011.
- ⁸ Zambia Health Management Information System Procedures Manual for Health Centers. Lusaka, Zambia: Central Board of Health, Directorate of Monitoring and Evaluation, 1998.
- ⁹ *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP, 2013.
- ¹⁰ Preston Samuel, Patrick Heuveline, Michel Guillot. *Demography: Measuring and modeling population processes*. Malden, MA: Blackwell Publishers, Inc: 2001
- ¹¹ Hill Kenneth, Alan D Lopez, Kenji Shibuya, Prabhat Jha. Interim measures for meeting needs for health sector data: births, deaths, and causes of death. *Lancet* 2007; 370: 1726–35.