Author’s response to reviews

Title: Errors in estimated gestational ages reduce the likelihood of health facility deliveries: results from an observational cohort study in Zanzibar

Authors:

Isabel Fulcher (isabelfulcher@g.harvard.edu)
Kaya Hedt (kayah@live.unc.edu)
Stella Marealle (smarealle@d-tree.org)
Jalia Tibajjuka (jaliatiba@d-tree.org)
Omar Abdalla (almuntaf@gmail.com)
Rachel Hofmann (rhofmann@d-tree.org)
Erica Layer (elayer@d-tree.org)
Marc Mitchell (mmitchel@hsph.harvard.edu)
Bethany Hdt-Gauthier (bethany_hedt@hms.harvard.edu)

Version: 1 Date: 20 May 2019

Author’s response to reviews:

Response to Reviewers

BHSR-D-19-00063: Errors in estimated gestational ages reduce the likelihood of health facility deliveries: results from an observational cohort study in Zanzibar

Mario Merialdi (Reviewer 1): This is a well written manuscript that suggests a provoking hypothesis that errors in EDD might influence the likelihood of delivery in a health facility. The authors correctly state that the data do not allow to draw definitive conclusions and indicate multiple factors that might confound the association as well as the limitations of their work. I think a factor that has not been fully considered is the potential contribution of biological variability to the observed variation in estimated delivery date. If possible it would be important to provide information of what could be the expected standard deviation of childbirth date estimated with a reliable method in this population or in a comparable one and relate this information to the intervals used to assess adequacy or under/over estimation of the EDD. In addition, it would be important to provide the biological rationale for choosing those intervals. Both questions could be probably addressed with the same data.
We thank the reviewer for appreciating the potential impact of our work. We agree that there is biological variability in the delivery date which we did not properly address this in the manuscript. A recent study showed that only ~5% of births occurred on their estimated day of delivery using reliable LMP and ultrasound dating (at various gestational ages) [2]. Further, the distributions of gestational ages indicated that about 91% of births occurred within +/- 14 days of the estimated date of delivery. The results were similar for all dating methods. However, we note that it is difficult to determine if this variability is driven by measurement error in the methods or natural variability in the population. We have mentioned this in the limitations with several citations:

Further, as this analysis only included women enrolled in the Safer Deliveries program in Zanzibar, Tanzania, the results may not be generalizable to other populations due to both regional differences in women’s accuracy in LMP recall, use of the Islamic vs. Gregorian calendar, and natural variability in gestational duration [2,11,13,27] and structural differences in maternal health programs (i.e. CHW experience, timing of visits, and delivery date estimation). (page 8, lines 31-36)

For creation of intervals, we now reference recent studies in both the US and LMICs that assessed the accuracy of gestational age dating by LMP. These studies considered tolerance windows of +/- 7 and +/- 14 days. Typically, &gt; 80% of deliveries were within 14 days of the estimated date. Another study found that LMP dating had a median absolute difference between estimated and actual of 7 days with an interquartile range of 2-12 days (references in text below). We have added the following to our main text directly following the definition of our categories:

The accurate category was defined based on multiple studies that reported actual delivery dates to be accurate within 7 and 14-days of the estimated delivery date due to LMP dating error and/or natural variability in length of gestation [1,2,13,14]. (page 5, lines 34-37)

José Fernando Georgia, MD, MPH (Reviewer 2): General Comments

The authors address a very important issue in developing countries with high maternal and infant mortality, such as Tanzania and neighboring countries.
Access to quality prenatal care is a key factor to address this challenge and trained community health workers (CHWs) are an important fill an important need as the country finds ways to expand access for prenatal care with trained health care providers such as trained midwives and physicians with appropriate obstetrical training.

This study in trying to address a key issue, gestational age as it is essential in planning the timing of visits during pregnancy at specific periods of gestational age.

The key aim is to investigate the magnitude and impact of errors in estimates delivery dates (EDD) on health facility delivery among women enrolled in the Safer Deliveries Program.

Categories of error in estimated delivery dates compared to actual delivery date (ADD) include (a) severe overestimate, defined as 36 days or more after actual delivery date; (b) moderate overestimate, when EDD was 15 to 35 days after ADD; (c) accurate, defines as 6 days before to 14 days after ADD; and (c) underestimate, defined as 7 days or more before ADD.

It is not clear to this reviewer the rationale for the definitions of overestimate, accurate, and underestimate. Studies that led to the adoption of best obstetrical estimate in the US showed that the LMP-based estimates were 0.8 days longer on average than ultrasound.


As we are interested in improvement of gestational age measurement (e.g. getting the estimated delivery date as close to the actual delivery date as possible), we tried to steer away from the classifications of preterm, term, and post-term. Thus, we looked at literature on estimation of gestational age by LMP to understand a reasonable window for error. Two studies in LMICs found that &gt;80% of deliveries were within 14 days of the estimated date (references in text below). To make clearer the reason for and creation of these categories, we have provided a more detailed explanation following our variable definition in the main text directly following the definition of our categories:
The accurate category was defined based on multiple studies that reported actual delivery dates to be accurate within 7 and 14-days of the estimated delivery date due to LMP dating error and/or natural variability in length of gestation [1,2,13,14]. Importantly, this variable captures the difference between when the woman expected to deliver and when that delivery actually occurred regardless of whether the difference was due to an error in estimated delivery dates (gestational age) or a true pre- or post-term birth. (page 5, lines 34-40)

Another consideration is that this analysis does not take into account preterm births as a reason for moderate to severe overestimate. How do the authors discriminate between error in the EDD estimate and short gestational age because of preterm delivery? For example, if a woman delivers a gestational age 34 weeks and the EDD was 6 weeks later, then EDD would fall within the correct estimate for term, but according to the categories, it would be considered severe overestimate. What this suggest that in addition to looking at errors in EDD, there is a need to identify risk factors for early delivery, such as preterm births.

We agree that we cannot distinguish between preterm or post-term births and error in the EDD estimate. However, the distribution of delivery dates (41% vs. the 11% from Tanzania) leads one to believe that most of these “preterm” births are likely errors in the EDD estimate. As the Reviewer correctly states, our analysis will include preterm births in the difference in delivery date categories.

In our study, we would have classified 42% of births as preterm if we believed the estimated delivery dates were correct. Although it has been reported that rates of preterm birth are highest in least developed regions, recent estimates of the preterm birth rates were only 14% in Eastern African countries and 10% in Tanzania [28,29,30]. Thus, it is likely that the estimated delivery dates recorded by the Safer Deliveries program are incorrect estimated delivery dates. Specifically, the severe and moderate overestimate categories include women who were truly preterm and those with an incorrect estimated delivery date. Due to the large discrepancy in gestational age distributions between the Safer Deliveries program and other studies conducted in this region, we believe the majority of the women in these categories have incorrect estimated delivery dates. (page 7, lines 6-15)
However, we emphasize that the inclusion of preterm births in such a way should not bias the findings from our main model; that is, we believe women who experience a preterm birth are not more (or less) likely to deliver in a health facility compared to women who deliver at term but with an incorrect estimated date of delivery (or gestation age).

Lastly, we were unable to classify if the births were truly preterm, term, or post-term due to the likely errors in gestational ages and unavailability of birth weight or physical assessments at birth in our data. This would only bias our findings if the decision to go to a health facility differed between women with an overestimated estimated delivery date and those who delivered preterm, which we do not believe is the case. (page 8, lines 37-41)

That being said, improvement in EDD will not apply to women who have preterm births. However, we note that the improvement of gestational age measurements in such maternal health programs will better allow for classification of preterm, term, and post-term deliveries. We have added the following discussion in the Discussion section:

Importantly, our findings also indicate that improvement of estimated delivery dates will not increase a woman’s likelihood of health facility delivery if she is truly preterm. However, as a byproduct of more accurate gestational age measurements, classification of preterm and post-term births will improve. This will enable the identification of preterm births at delivery such that neonates can receive appropriate and timely medical care [29,30]. In addition, the collection of quality data on preterm birth occurrences will enable assessment of the potential risk factors within a population, which can then be used to predict future preterm births in a population. (page 8, lines 17-24)

The authors need to explain the rationale for their categories of error and why preterm birth was not considered in this analysis.

As described previously, we have now added rationale for the categories for error. In addition, we provide a thorough discussion regarding preterm birth classifications at multiple points in the discussion section.
Another aspect to consider to describe in the methods is what is the general approach for clinically recognizing a pregnancy? Is it based on missing two consecutive periods, a pregnancy test, clinical evidence of an enlarged uterus? This is an important aspect to describe.

When pregnant women attend their first ANC clinic, they are provided with a pregnancy test to confirm. In the absence of reagent for pregnant test, the clinical method used is missing two consecutive periods and clinical evidence of an enlarged uterus. If an ultrasound is available, they may also use this to determine pregnancy. For women who present for an ANC visit later in their pregnancy (after 7 months), a test may not need to be performed. We have added in the following statement:

Typically, a pregnancy is confirmed at the health facility during the first antenatal care visit using a pregnancy test. In the absence of reagent for the pregnant test, missing two consecutive periods and clinical evidence of an enlarged uterus is used to determine pregnancy status. (page 4, lines 19-22)

The discussion should consider the issue of preterm births and the potential options of educating women about the early warning signs of a preterm births and the importance of seeking medical assistance at a medical facility.

We thank the reviewer for this interesting suggestion. We have added the following in the statement in our Discussion section:

However, as a byproduct of improved gestational age measurements, preterm and post-term births will be more accurately identified. This will ensure that preterm neonates receive appropriate and timely medical care [29,30]. In addition, the collection of quality data on preterm birth occurrences will enable assessment of the potential risk factors within a population, which can then be used to predict future preterm births in a population. (page 8, lines 18-24)

Specific Comments

Page 3, line 33-38.
This sentence misses a very crucial point and is that there is great variation worldwide in the rate of preterm births as well as of post-term births. Suggest that authors consider the following references:

Lawn et al. BMC Pregnancy and Childbirth 2010, 10(Suppl 1):S1
http://www.biomedcentral.com/1471-2393/10/S1/S1
https://www.who.int/pmnch/media/news/2012/201204_borntoosoon-report.pdf

Both reference provide more comprehensive data on the distribution of preterm births around the world.

As the goal of this paper is to discuss errors in estimated delivery dates due to LMP, we only referenced studies that were investigating the accuracy of LMP in similar settings. This is not necessarily the case for the estimates in these more comprehensive papers, which seem to use a combination of best obstetric estimates, ultrasounds, LMP, and other unspecified methods for gestational age measurement. We have tried to make this distinction clear by modifying the following statement in the Introduction:

Importantly, due to differences in study designs and populations, these distributions are not directly comparable, but provide a range of plausible birth distributions from LMP measurements covering a variety of settings. (page 3, lines 25-27)

In addition, we have moved the Reviewer’s suggested modification and valuable references to the Discussion section:
In our study, we would have classified 42% of births as preterm if we believed the estimated delivery dates were correct. Although it has been reported that rates of preterm birth are highest in least developed regions, recent estimates of the preterm birth rates were only 14% in Eastern African countries and 10% in Tanzania [28,29,30]. Thus, it is likely that the estimated delivery dates recorded by the Safer Deliveries program are incorrect estimated delivery dates. (page 7, lines 6-10)

The authors state the data on whether the EDD was based on LMP or ultrasound and then state that it was assumed that most were based on LMP and the assumption is made that all were LMP-based. On the other hand, the authors also state that Safer Deliveries was active in six out 11 districts Zanzibar on the islands if Unguja and Pemba. This reviewer wonders if the authors can at least determine which of the six areas had access to ultrasound and which did not and look at the distribution of categories of errors on EDD between those sites with access to ultrasound and those who did not.

We greatly appreciate the reviewer’s suggestion and were able to determine that two Primary Health Care Centers (PHCC) and two hospitals in Zanzibar had access to ultrasound machines during the time of our study. Using information on location of antenatal care visits, 419 (10%) of women attended at least 1 ANC visit at these locations, but we have no way of determining whether these women actually received an ultrasound. There was no appreciable difference in the distribution of preterm, term, or post-term classified births at these locations.

Figure. Distribution of birth term classification* by ultrasound availability at ANC site

Ultrasound not available (n=3,785)

Pre-term: 1,579 (41.7%)

Term: 2,033 (53.7%)

Post-term: 173 (4.6%)
Ultrasound available (n=419)

Pre-term: 181 (43.2%)
Term: 220 (52.5%)
Post-term: 18 (4.3%)

P-value (Chi-squared test): 0.834

Note: 21 (<1%) women have missing values for ANC visit locations

* If term of the neonate is classified based on the date of birth relative to the estimated delivery date

We have now updated our statement regarding ultrasounds in our Methods section:

Although we did not collect information on whether EDD was ascertained by date of LMP or ultrasound, only two Primary Health Care Centers and two hospitals in Unguja have an ultrasound machine available which may not even be commonly used. Further, only 419 (10%) women reported an ANC visit at one of these health centers, and the distribution of preterm, term, and post-term classifications did not significantly differ from women without ultrasound access at their ANC visit(s) (P = 0.834). Due to this, we believe that the vast majority of the delivery dates were calculated based on LMP at an antenatal care or community health worker visit. (page 5, lines 10-17)

We have also removed the limitation about potential confounding by method of estimation (ultrasound vs. LMP) in the Discussion section as this no longer is a relevant issue.
Categories of error in EDD estimate. There is no rationale described for the selected categories of severe underestimate and moderate underestimate. Moreover, there is apparent consideration to the fact of preterm births, something that can be assessed with birth weight and physical assessment. Not considering preterm births is a major flaw of this categories since it does not discriminate between the error in estimating EDD and the early delivery resulting in a preterm births. This becomes evident in figure 1 that reports 42% as preterm, which suggest that this estimate is a combination of error in EDD and preterm births. According to UNICEF and the Global report on Preterm births the preterm birth in Tanzania in 2010 was 11%, quite different for the 42% estimated in this study from EDD calculated from LMP.

https://data.unicef.org/wp-content/uploads/country_profiles/United%20Republic%20of%20Tanzania/country%20profile_TZA.pdf

We thank the reviewer for this Tanzania specific reference and have added it on page 7, line 3. As previously described, we now include rationale for our category choices in the text and a discussion on preterm births at multiple points throughout the paper. While we agree that the ability to distinguish preterm births is imperative for quality care, we do not believe that the presence of preterm (or post-term) births in this study biases the results from our main model. Specifically, we believe that the decision to go to a health facility does not differ between women with an overestimated estimated delivery date and those who are delivering preterm. We have added in the following statement:

Lastly, we were unable to classify if the births were truly preterm, term, or post-term due to the likely errors in gestational ages and unavailability of birth weight or physical assessments at birth in our data. This would only bias our findings if the decision to go to a health facility differed between women with an overestimated estimated delivery date and those who delivered preterm, which we do not believe is the case. (page 8, lines 36-41)

That being said, we agree that our final conclusions do not apply to women who deliver preterm. That is, improvements in EDD measurements will not improve their probability of health facility delivery unless additional action is taken. As previously mentioned, we have added detail discussion on this point:
Importantly, our findings also indicate that improvement of estimated delivery dates will not increase a woman’s likelihood of health facility delivery if she is truly preterm. However, as a byproduct of more accurate gestational age measurements, classification of preterm and post-term births will improve. This will enable the identification of preterm births at delivery such that neonates can receive appropriate and timely medical care [29,30]. In addition, the collection of quality data on preterm birth occurrences will enable assessment of the potential risk factors within a population, which can then be used to predict future preterm births in a population. (page 8, lines 17-24)

Page 6, line 12-16. To report that the preterm rate was 42% based on difference of LMP and date of birth may not be accurate without an independent assessment of birth weight and gestational age based on physical exam of the newborn.

We do not believe that the preterm rate was actually 42%, and we realize that this was not clear in the paper. In fact, we believe that the majority of births that were classified as preterm are actually just errors in estimated delivery dates. We have added in the following statements to clarify:

In the Safer Deliveries data, if we classified term of the neonate based on the date of birth relative to the estimated delivery date, then 42% of births would be classified as preterm (< 37 weeks), 50% term (37-40 weeks), and 8% post-term (> 41 weeks). (page 6, lines 19-21)

In our study, we would have classified 42% of births as preterm if we believed the estimated delivery dates were correct. (page 7, lines 6-7)

Page 7: lines 54-60 Should include in the limitations the fact that preterm births was not considered in the assessment of errors in EDD.

As stated previously, we have added the following into our limitations paragraph:
Lastly, we were unable to classify if the births were truly preterm, term, or post-term due to the likely errors in gestational ages and unavailability of birth weight or physical assessments at birth in our data. This would only bias our findings if the decision to go to a health facility differed between women with an overestimated estimated delivery date and those who delivered preterm, which we do not believe is the case. (page 8, lines 36-41)