Community workers contribution for data quality improvement in routine health information system: Live births number estimation by capture recapture method

Yolaine Glèlè Ahanhanzo1,2*, Alphonse Kpozéhouen1, Noël Moussiliou Paraïso3, Alain Levêque2, Michel Makoutodé4, Michèle Dramaix-Wilmet2

1Epidemiology and Biostatistics Department, Public Health Regional Institute, Ouidah, Benin
2Center of Research in Epidemiology, Biostatistics and Clinical Research, School of Public Health, Université Libre de Bruxelles, Brussels, Belgium
3Health Promotion Department, Public Health Regional Institute, Ouidah, Benin
4Environment and Health Department, Public Health Regional Institute, Ouidah, Benin
Email: *nyglele@yahoo.fr, nygleleah@ulb.ac.be

Received 9 January 2014; revised 9 February 2014; accepted 16 February 2014

Copyright © 2014 Yolaine Glèlè Ahanhanzo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for SCIRP and the owner of the intellectual property Yolaine Glèlè Ahanhanzo et al. All Copyright © 2014 are guarded by law and by SCIRP as a guardian.

ABSTRACT

The lack of reliable vital statistics raises questions about the role of the health information system in acquiring such data, which are essential for planning health services and for the general management of the population’s needs. This study analyzed completeness of the vital data registration system and assessed the potential contribution of a community worker network to this system in rural Benin. The capture-recapture method was used in this interventional study to estimate the number of live births from three sources: the Routine Health Information System, the municipality, and community workers in two groups of villages. Log linear modelling was carried out with a Bayesian Information Criterion-weighted estimate of the number of live births. The exhaustiveness of the Routine Health Information System was improved by the contribution of the community workers from 29.3% to 42.5% in the first group, and from 61.7% to 77.5% in the second group. Estimating live births by the capture method in rural settings based on the contribution of community workers could be a more efficient alternative to censuses in acquiring reliable vital statistics.

KEYWORDS
Community Health Workers; Health Information System; Vital Statistics; Capture-Recapture Method; Benin

1. INTRODUCTION

To ensure the population’s health, and to efficiently plan for healthcare needs, governments must have comprehensive and reliable demographic data. In this respect, vital statistics are essential. According to the United Nations, the best source is the civil registration system for the continuous, permanent and compulsory registration of the vital events in a person’s life, such as live births and deaths, in accordance with the standards of each country [1]. Health Information Systems (HIS) play a key role in providing this information due to the fact that they are best placed in relation to vital events such as births and deaths. More particularly, Routine Health Information Systems (RHIS), information derived at regular intervals from mechanisms designed to meet predictive information needs, should be capable of providing this information [2,3]. The role of RHIS, in some countries, is even more significant given the current challenge of achieving the MDGs. In fact, reliable information concerning maternal and infant mortality is needed to be able to assess the level of achievement of the MDGs [2,4].

While in developed countries HIS provide most of this information, the low performance of HIS in developing countries does not make it possible to obtain this vital data [1,5-8]. Although they are more or less operational in most developing countries, HIS are currently described as ineffective for several reasons, such as the insufficient exhaustiveness, quality and use of the data collected [9-12].
In Benin, the RHIS, known as the Système National d’Information et de Gestion Sanitaire (SNIGS), is based on data routinely collected from health facilities. It comprises several sub-systems, including the birth registration sub-system. Various internal and external evaluations of the SNIGS agree on the poor performance of this sub-system [9,13,14]. According to the World Health Statistics 2013, birth registration coverage in Benin from 2005 to 2011 was 60% [15].

To improve the exhaustiveness of information concerning vital events, various strategies have been proposed and implemented at national level in countries such as Ghana and India, with varying degrees of success [6,8,16]. On a more local level, in Malawi, Indonesia and DRC, studies have identified positive experiences with community-based organisations [4,17-19]. According to these studies, the workers of community-based organisations or community workers (CW) could represent an opportunity to improve the performance of RHIS as regards the exhaustiveness of vital events.

Community workers play a key role in health systems in developing countries. Many projects are implemented at community level thanks to these players whose integration within the health system and motivation remain unresolved [20,21]. In some health zones in Benin, there are a number of community worker networks more or less operational under community projects financed by technical and financial partners. This study aimed to assess the contribution of an existing operational community worker network to improving the performance of RHIS in reporting live births in rural Benin.

2. CONTEXT AND STUDY METHODS

2.1. Context

The study was conducted in the health zone of Ouidah in the south of Benin. The health zone covers an essentially rural population of approximately 300,000 inhabitants. The strategy of providing community care for malaria cases, with the support of technical and financial partners, has been in application in some of the villages in this health zone since 2008. This strategy mainly relies on community worker networks trained for the purposes of the project. These community workers therefore have close contact with this project’s target groups, which are children under five and their mothers in each village. Each village involved in community-based malaria care has at least one trained community worker. The community workers work under the supervision of the health worker responsible for the health centre that covers the village. They provide a monthly activity report which gives an overview of all activities concerning treatment, referrals and medicine transfers.

2.2. Methods

2.2.1. Type of Study

This study was an interventional one and analysed the exhaustiveness of the vital data registration system using the capture-recapture method [22-25].

2.2.2. Village Samples

Fourteen (14) villages were selected from within the health zone based on the effective level of operation of the CW network and its activities, along with the expected stability of the populations in these areas. Villages accommodating nomadic populations were excluded. The level of operation was assessed on the availability of activity reports over the last six months, and the final selection was made with the district medical officer and the health workers overseeing the project. The villages were divided into two groups for the collection of data. Group 1 comprised 8 villages and group 2 was made up of 6 villages. According to estimates made on the basis of the last population census, 170 live births were expected in group 1 and 77 live births were expected in group 2 for the period covered by the study; the rate of health service use for childbirth was 43.2% in group 1 and 80.3% in group 2 [26]. In group 1, the community workers collected information relating to live births in their village retrospectively and, in group 2, this data collection was a forecast.

2.2.3. Data Sources

In this study, three sources of data were used: the community workers through a community survey, the RHIS, and the municipalities through the district or municipal birth registry.

Community Survey: The community workers of each village were invited to produce a list of all the live births that they knew of and that had occurred in their village during the three months preceding the day of the survey. On the basis of this list, agreed and validated by the community worker or contacts and supplemented with information needed to identify the households concerned, a community survey was conducted among the households to confirm whether or not a live birth had actually taken place in the period under analysis and to collect the associated data. In the villages from group 1, the community workers were not warned in advance of the survey, so the information was collected by the community workers retrospectively, from their memory and from information from their local network. In the villages in group 2, a working session with the community workers was held four months before the survey. This session clarified the results expected from the community workers and emphasised the passive nature of the data collection carried out by them; they were to avoid seeking the
information from other sources. Unlike the first group, the collection of information by the community workers in this second group of villages was a forecast. This approach in the second group aimed at improving the exhaustiveness and quality of the information provided by the community workers and at better ensuring that the community workers and the other sources of information were independent of one another.

Routine Health Information System (RHIS): The RHIS data collection tool, the birth records available at health centre level, provides all the information relating to deliveries and births in the public health centres and village health units, and in certain private centres participating in the RHIS.

District or Municipal Births Registry: The registry compiles all the information concerning births that have taken place within the geographical area of the municipality. All births must be registered within ten days. When the birth occurs in a public or legal private health facility, the parents are issued with a birth record so that they can register the birth with the municipality. When the birth takes place outside of a public or legal private health facility, the parents must go directly to the municipality to register the birth. The municipality is, in theory, the compulsory declaration source, but there is no penalty if the child is not declared and, at present, parents register births if they want to.

2.2.4. Survey Process
The data collection was carried out by three groups of interviewers trained separately and sent separately to the three sources of data to identify live births. The household survey and the data collection from the health centre RHIS were carried out at the same time while the data collection from the birth registries was carried out 15 days later.

2.2.5. Operational Definition of Live Birth
We used the WHO definition, cited by Statistics Netherlands Office, according to which a live birth is a birth with any other evidence of life [27].

2.2.6. Cross-Referencing Variables
The variables collected and used to cross-reference the data were: the date of birth, the child’s sex, the father’s name and surname, the mother’s name and surname, and the place of residence.

2.2.7. Data Analysis and Processing
The data was recorded in Excel 2007 and processed with Stata 13. The database was cleaned to remove duplicates. Multiple births were taken into account once the sources had been confirmed. We evaluated the dependency of the sources on each other with odds ratios (OR) and with their confidence interval at 95% (CI95%) [28,29]. We carried out log linear modelling. All possible combinations were tested ranging from the model with three independent sources to the saturated model with three sets of interaction between pairs of sources. The selection of plausible models was based on the lowest Akaike Information Criterion (AIC) and Draper’s Bayesian Information Criterion (BIC) values [30,31]. For all of the models, we estimated the number of births not reported by any of the three sources. From this, we deduced an estimate of the total number of births for each model and an average BIC-weighted estimate of the total number of births on the basis of the models selected [22,30,31]. A corrective adjustment was made (add 1) for cells with very low values [25,32].

2.2.8. Ethical Considerations
The research protocol for this project was approved by the Benin National ethics committee for research in health (Comité National d’Ethique pour la Recherche en Santé) and also by the regional authorities and the local health authorities. The community workers and the household respondents were given an information form and gave their informed consent before participating in the study.

3. RESULTS
In total, 391 live births were reported in group 1 and 98 in group 2 for the period covered by the study. These were distributed between the groups of villages as shown in Figures 1 and 2.

In group 1, half of the 81 births not identified by the municipality were reported by each of the other sources and there was a positive dependency between the sources taken in pairs. As such, \( OR_{RHS-CW} = 4.2; \text{ CI95%} = [2.3 - 7.8] \), \( OR_{Municipality-CW} = 8.9; \text{ CI95%} = [2.0 - 39.3] \), \( OR_{RHS-Municipality} = 31.3; \text{ CI95%} = [6.8 - 143.2] \). \( OR_{RHS-CW} \) was the lowest; the dependency between the CW source and the RHIS was less significant. In group 2, 24 births were not identified by the municipality and there was no de-
Dependency between the CW source and the RHIS source (OR_{RHIS-CW} = 0.51; CI_{95\%}: [0.2 - 1.6]). There was a positive dependency between the other sources taken in pairs. As such, OR_{RHIS-Municipality} = 7.4; CI_{95\%}: [1.6 - 135.0].

![Figure 2. Distribution of live births according to the sources of information in group 2 villages.](image)

**Table 1. Estimation of live births number (N) and weighted mean of live births number (NBIC) in group 1.**

<table>
<thead>
<tr>
<th>Models</th>
<th>df</th>
<th>G²</th>
<th>AIC</th>
<th>BIC</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saturated model: CW, RHIS, Municipality, CW * RHIS, CW * Municipality, RHIS * Municipality</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>478</td>
</tr>
<tr>
<td>Model 2: CW, RHIS, Municipality, CW * Municipality, RHIS * Municipality</td>
<td>1</td>
<td>21.7</td>
<td>19.7</td>
<td>20.7</td>
<td>1170</td>
</tr>
<tr>
<td>Model 3: CW, RHIS, Municipality, CW * RHIS, RHIS * Municipality</td>
<td>1</td>
<td>14.0</td>
<td>12.0</td>
<td>13.0</td>
<td>760</td>
</tr>
<tr>
<td>Model 4: CW, RHIS, Municipality, CW * RHIS, CW * Municipality</td>
<td>1</td>
<td>37.4</td>
<td>35.4</td>
<td>36.5</td>
<td>497</td>
</tr>
<tr>
<td>Model 5: CW, RHIS, Municipality, RHIS * Municipality</td>
<td>2</td>
<td>26.6</td>
<td>22.6</td>
<td>24.7</td>
<td>611</td>
</tr>
<tr>
<td>Model 6: CW, RHIS, Municipality, CW * Municipality</td>
<td>2</td>
<td>40.1</td>
<td>36.1</td>
<td>38.2</td>
<td>485</td>
</tr>
<tr>
<td>Model 7: CW, RHIS, Municipality, CW * RHIS</td>
<td>2</td>
<td>39.1</td>
<td>35.1</td>
<td>37.2</td>
<td>516</td>
</tr>
<tr>
<td>Independent model: CW, RHIS, Municipality</td>
<td>3</td>
<td>42.8</td>
<td>36.8</td>
<td>40.0</td>
<td>505</td>
</tr>
<tr>
<td>Weighted mean NBIC</td>
<td>478</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Selected models; *Weighted mean of live births number for selected models; G²: likelihood ratio statistic; df: degrees of freedom, AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion.

**Table 2. Estimation of live births number (N) and weighted mean of live births number (NBIC) in group 2.**

<table>
<thead>
<tr>
<th>Models</th>
<th>df</th>
<th>G²</th>
<th>AIC</th>
<th>BIC</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saturated model: CW, RHIS, Municipality, CW * RHIS, CW * Municipality, RHIS * Municipality</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Model 2: CW, RHIS, Municipality, CW * Municipality, RHIS * Municipality</td>
<td>1</td>
<td>1.3</td>
<td>-0.7</td>
<td>0.6</td>
<td>149</td>
</tr>
<tr>
<td>Model 3: CW, RHIS, Municipality, CW * RHIS, RHIS * Municipality</td>
<td>1</td>
<td>17.5</td>
<td>15.5</td>
<td>16.8</td>
<td>100</td>
</tr>
<tr>
<td>Model 4: CW, RHIS, Municipality, CW<em>RHIS, CW</em>Municipality</td>
<td>1</td>
<td>8.4</td>
<td>6.4</td>
<td>7.7</td>
<td>102</td>
</tr>
<tr>
<td>Model 5: CW, RHIS, Municipality, RHIS * Municipality</td>
<td>2</td>
<td>22.1</td>
<td>18.1</td>
<td>20.6</td>
<td>104</td>
</tr>
<tr>
<td>Model 6: CW, RHIS, Municipality, CW * Municipality</td>
<td>2</td>
<td>11.9</td>
<td>7.9</td>
<td>10.5</td>
<td>105</td>
</tr>
<tr>
<td>Model 7: CW, RHIS, Municipality, CW * RHIS</td>
<td>2</td>
<td>17.9</td>
<td>13.9</td>
<td>16.4</td>
<td>100</td>
</tr>
<tr>
<td>Independent model: CW, RHIS, Municipality</td>
<td>3</td>
<td>23.7</td>
<td>17.7</td>
<td>21.5</td>
<td>102</td>
</tr>
<tr>
<td>Weighted mean NBIC</td>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Selected models; *Weighted mean of live births number for selected models; G²: likelihood ratio statistic; df: degrees of freedom, AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion.
4. DISCUSSION

The capture-recapture method is used to assess the exhaustiveness of information between different sources. It has been applied in a number of epidemiological studies [18,33-35]. It makes it possible, when presented with several sources of information for the same population or the same phenomenon, to estimate the number of cases not identified by any of the sources, to extrapolate the actual total number of cases of the phenomenon being studied and to determine the exhaustiveness of each of the sources based on this estimated total number of cases. This method seems appropriate to answer the questions raised in this study as long as a number of conditions are taken into consideration.

Case definition does not pose a problem concerning the identification of real cases; indeed, identifying false cases for live births is highly improbable. However, despite all the precautions taken during the community survey, and given the sensitive nature of the issue of death, it is possible that there was, to a lesser extent, under reporting of cases by the community sources and the municipality due to the possible non-registration of neonatal deaths which are in fact live births.

The period and geographic zone were adhered to as much as possible in this study as all the cases taken into account were residing in the villages covered by the study. We took the period aspect into account and ensured that the collection of information at municipal level took the legal deadline for registering births with the regional authorities into account.

The nature of the event under analysis (live births) could seem to have an impact on the closed population condition. Nevertheless, one of the criteria for selecting the villages was also the stability of the populations in order to meet this condition. Moreover, as highlighted by Schmid and Silva in their study in Brazil in 2011, even if the parents move in the meantime, this does not change the fact that the live birth took place within the area covered by the study [34].

The combination of several cross-referencing variables (five cross-referencing variables) to identify cases enabled us to better ensure that all of the real common cases, and only the real common cases, were identified. Compliance with this condition may in fact be questionable when the identification is based on a single condition like the surname, which can be spelled several ways, or in a situation where the pronunciation or the interviewer’s phonetic understanding can influence the identification [33,36]. In the context studied, as the lifestyles and behaviours are relatively stable within the population, the probability of capturing live births is not heterogeneous within the same source.

In our study, and as regards the two groups, the analysis of the sources taken in pairs revealed a positive dependency, which points to the likelihood of an underestimation in the total number of cases. This dependency is easily explained given the many awareness raising campaigns conducted in the country, in connection with the health system, for the registration of births. Some municipalities even require a birth record from a health facility even if the birth did not take place in a health centre. The significant difference observed in the two groups concerning the dependency between RHIS and the municipality could be explained by the customs of the municipalities and the differences in birth registration procedures. Moreover, the CWs are leading people in their social environment. In general, they are figures who, given their position in the village, may also be involved in these awareness campaigns. However, the normal and official process for reporting births does not involve the CWs and there is normally no dependency between the CW and RHIS sources; this was demonstrated in this study in group 2 where preparatory work was carried out with the CWs to clarify the results expected. The difference observed with group 1 could be explained by the interaction between the various CWs to validate the list which served as a basis to the household survey or by the various interactions that took place between the CWs and the health centres for other activities. The dependency between sources is common and can lead to the grouping of 2 sources [28,29], but it remains difficult to completely dismiss in a study where there is interaction between the members of the same community [18].

The number of live births observed was greater than the number of births expected according to national statistics [26]. It is clear therefore that there is a significant underestimation in the number of births and this raises questions about the reliability of the methods used for these estimates, such as censuses and extrapolations based on birth rates. Based on the different corrective adjustments used to take into account the dependency between the sources, in group 2, the exhaustiveness of the municipal source—a compulsory birth declaration source—was close to WHO estimates which were 60% [15]. The exhaustiveness of the RHIS was 29.3% in group 1 and 61.7% in group 2. This variation could be explained by the difference in health service use between the 2 groups [26].

In group 2, the exhaustiveness of the CW source is significant (about two-thirds of the legal system) and it was better than in group 1; in theory, prospective data collection should improve their contribution. In group 1, although the exhaustiveness of the CW source was lower than in group 2, the CWs alone identified as many cases not declared in the legal system as the RHIS, if not more (43 of the 81 cases not identified by the municipality). In our study, the number of live births that had not been
reported by any sources amounted to about 18% which is similar to the results of the study to estimate the number of unreported births in the north east of Brazil [34], but is higher than the 9% reported in another Brazilian study in 2008 [37]. This difference could be explained by the fact that the latter was based solely on hospital facility health information systems. In fact, the hospital HIS is not representative of births in developing countries because more often the deliveries that occur in hospital facilities are at-risk pregnancies or referred parturient cases. If we do not take into account the CWs in our study, the number of live births not reported by any source rises from 18.2% to 26.8% in group 1 and from 18.3% to 23.3% in group 2. These results bear out the positive benefit of the CWs in improving the vital statistics system, as reported by studies in Malawi, Indonesia and in DRC [17-19]. The estimates made in group 1 are questionable as the optimum model seems to be the saturated model [23,31]. However, these estimates remain interesting as they are realistic and similar to figures reported in literature concerning the exhaustiveness of the municipal source [15]. They also highlight, in this study, the benefit of keeping the CW source independent of the RHIS and municipality sources in terms of implementation strategy in order to have an accurate estimate of the number of live births. This study was deliberately conducted on the basis of a CW network that was already in operation. The investigation did not lead to active research by the network for two reasons: i) estimating the potential passive contribution did not lead to active research by the network for a CW network that was already in operation. The investigation could have an accurate estimate of the number of live births. ii) avoiding raising additional but very relevant questions concerning the motivation of these players. These issues could not be avoided and should be taken into account in terms of project strategy in order to improve the performance of the RHIS based on the contribution of the CWs. Taking the CWs into consideration in the reporting and registration system for live births will no doubt help improve the exhaustiveness of the system. To have optimum estimates, in terms of strategy, the CWs must be integrated as a source of additional, supplementary information and not be directly integrated in the birth reporting process in order to guarantee their independence in relation to the more traditional sources of information. Such a strategy will most certainly be worthwhile in terms of cost reductions as it is probably more efficient than censuses which do not provide any more reliable results. Implementing such a strategy should nevertheless include conditions concerning the method’s application, particularly as concerns the independence of the CW source in relation to more traditional sources in order to have optimum estimates of vital events.

5. CONCLUSION

In a setting such as Benin, where reliable vital statistics are not available, the capture-recapture method remains a valuable approach for estimating the number of live births in rural settings. This method revealed, in this study, the positive contribution of the community workers in improving the exhaustiveness of RHIS data. The development of a strategy to estimate the number of live births based on the contribution of the CWs may prove more efficient than censuses and major surveys which do not provide any more reliable results. Implementing such a strategy should nevertheless include conditions concerning the method’s application, particularly as concerns the independence of the CW source in relation to more traditional sources in order to have optimum estimates of vital events.

REFERENCES


