RESEARCH ARTICLE

# Improved calibration estimators for the total cost of health programs and application to immunization in Brazil

**Claudia Rivera-Rodriguez**[1]*, **Cristiana Toscano**[2], **Stephen Resch**[3]

**1** Department of Statistics, The University of Auckland, Auckland, New Zealand, **2** Department of Community Health, Institute of Tropical Pathology and Public Health, Federal University of Goiás (UFG), Goiás, Brazil, **3** Center for Health Decision Science, Harvard T.H. Chan School of Public Health, Boston, MA, United States of America

* c.rodriguez@auckland.ac.nz, clriverarodriguez@gmail.com

## Abstract

Multi-stage/level sampling designs have been widely used by survey statisticians as a means of obtaining reliable and efficient estimates at a reasonable implementation cost. This method has been particularly useful in National country-wide surveys to assess the costs of delivering public health programs, which are generally originated in different levels of service management and delivery. Unbiased and efficient estimates of costs are essential to adequately allocate resources and inform policy and planning. In recent years, the global health community has become increasingly interested in estimating the costs of immunization programs. In such programs, part of the cost correspond to vaccines and it is in most countries procured at the central level, while the rest of the costs are incurred in states, municipalities and health facilities, respectively. As such, total program cost is a result of adding these costs, and its variance should account for the relation between the totals at the different levels. An additional challenge is the missing information at the various levels. A variety of methods have been developed to compensate for this missing data. Weighting adjustments are often used to make the estimates consistent with readily-available information. For estimation of total program costs this implies adjusting the estimates at each level to comply with the characteristics of the country. In 2014, A National study to estimate the costs of the Brazilian National Immunization Program was initiated, requested by the Ministry of Health and with the support of international partners. We formulate a quick and useful way to compute the variance and deal with missing values at the various levels. Our approach involves calibrating the weights at each level using additional readily-available information such as the total number of doses administered. Taking the Brazilian immunization costing study as an example, this approach results in substantial gains in both efficiency and precision of the cost estimate.

## Introduction

The global health community has become increasingly interested in estimating the costs of delivering public health programs, which are generally originated in different levels of service management and delivery. The interest has arisen from the need for financial sustainability and expansion of these programs. However, in low and middle income countries the resource utilization and financial information is scarce. Among the reasons for the lack of such information is the fact that public health program implementation occurs mostly at decentralized levels of management, and many costs are shared with other programs. As a result, program managers do not have the key information to improve efficiency or to advocate for adequate budget support. This is even more important when considering newly available technologies that are incorporated into public health programs, such as the introduction of new vaccines which are significantly more costly that the traditional childhood vaccines [1–4].

The costs of immunization programs occurs in various levels, where part of the cost correspond to vaccines and it is in most countries procured at the central level, while the rest of the costs are incurred in states, municipalities and health facilities, respectively. As such, total program cost is a result of adding these costs, and its variance should not only account for the uncertainty at each level, but also for the uncertainty due to the totals at each level being related (or nested) and for missing data.

In the past several years, selected studies have been conducted to assess the costs of immunization programs. In particular, various low and middle income countries conducted immunization costing studies since 2010, with the support from the Expanded Program on Immunization Costing and Financing (EPIC) project [1]. Detailed costing information on routine child immunization from Benin, Ghana, Honduras, Moldova, Uganda and Zambia [2, 3, 5–8]. These studies had in common the fact that they aimed at estimating total immunization costs, and not only expenditures on immunization. They were characterized by using multi-stage sampling of facilities and regression modelling to estimate the average costs at the various levels. The estimated average cost was then multiplied by the nationwide number of corresponding units (e.g. districts, facilities, etc.) to estimate total program costs. The study conducted in Honduras used sampling weights to obtain estimates of total national costs for routine immunization [3]. Despite having many strengths, the methods require that information is not missing, and does not allow generating a measure of uncertainty for the estimated total immunization program cost.

Over recent years, a variety of methods have been developed to compensate for missing data [9–17]. Weighting adjustments are often used to make the estimates consistent with readily-available information. For estimation of total program costs this implies adjusting the estimates at each level to comply with the characteristics of the country. Due to the need for collecting data at various levels, the large number of health-care units providing immunization at the local level, and considering the wide diversity of cost categories in immunization programs, it is inevitable that selected data will be lacking when conducting a comprehensive costing survey of immunization programs. Such missing data may be from the sampling units at any of the levels surveyed. This, coupled with the lack of an uncertainty measure of the overall cost estimate, represents the major methodological challenges when conducting nationwide immunization costing studies. Obtaining a precise and accurate cost estimator is critical for decision makers to inform policy and planning, and to recommend future studies.

Brazil structured its National Immunization Program in the 70's, and has since strengthened the program with legislation to secure funds. Various newly available vaccines have been incorporated into the program in the past decade, taking into consideration recommended framework for vaccine introduction decision-making [4]. In order to generate evidence to

support decision making and strengthen national capacity to appropriately inform policy and planning on the introduction of new vaccines, the Brazilian ministry of health commissioned a study to assess the costs of the Brazilian National immunization Program. A nationwide costing survey was conducted considering the year of 2013 and the perspective of SUS, the Brazilian National Public Health-care system which provides immunization services free of charge to all population through its more than 35,000 health-care services distributed in the country's 5 macro-regions and 27 states (including a federal district) [4]. More details of the methods can be found in [6].

In brief, data on all resources used in the immunization activities which occur at four organizational levels was collected: central, state, municipal, and health care facilities providing immunization services. Data was collected at the central and all 27 state levels, in addition to a sample of municipalities and health-care facilities. The sample was collected following a two-stage stratified scheme, where the strata and stages are based on regions and municipalities respectively. After data collection, selected data was missing at both the facility and municipality levels.

Using the Brazilian National Immunization Program Costing Study, we propose and validate a quick and useful approach to estimating the inter-level variance while accounting for missing data at different sampling levels. Our method involves calibrating the weights at each level using additional readily-available information for the corresponding entire level such as the total number of doses administered or population targeted by the immunization program. Furthermore, we present theoretical calculations on how to compute the variance of this estimator using Taylor approximations. The paper is composed as follows. The first section describes the total cost and its components, the second section describes the sampling design and missing data, followed by estimation methods and variance. This is followed by post-design methods to improve estimation of the total cost and we finalize the paper with application to the Brazilian immunization study.

## The total cost $T$

The main interest lies in estimating the total cost of a health program, e.g. the routine immunization program in Brazil. This cost can be expressed as

$$T = T_{\mathrm{C}} + T_{\mathrm{F}} + T_{\mathrm{M}} + T_{\mathrm{R}}, \tag{1}$$

where $T_{\mathrm{F}}$, $T_{\mathrm{M}}$ and $T_{\mathrm{R}}$ represent the total costs at the facility level, municipality level and regional level, respectively. Note that $T_{\mathrm{F}}$, $T_{\mathrm{M}}$, or $T_{\mathrm{R}}$ are not necessarily unknown, but it is usually the case that at least two of them are to be estimated [2, 3, 5, 7, 8]. $T_{\mathrm{C}}$ represents the additional cost accrued centrally and it is assumed known. For example, in Brazil, this cost arises from vaccines purchased at the central level, in addition to coordination and management activities incurring in personnel, transportation and infrastructure costs. Note that, if desired, this expression can also be used to estimate any total figure of interest. Usually costing studies can have different goals, but, in order to simplify and facilitate the readability of the paper, we focus on total cost. We denote $N_{\mathrm{R}}$ as the total number of regions, $N_{M_r}$ as the total number of municipalities within region $r$ and $N_{F_r m}$ is the total number of facilities in municipality $m$, region $r$. Each of these costs can, respectively, be expressed as follows:

$$T_{\mathrm{R}} = \sum_{r=1}^{N_{\mathrm{R}}} y_{R_r}, \tag{2}$$

where $y_{R_r}$ is the total cost specific to regions and $N_{\mathrm{R}}$ is the number of regions. Similarly, $T_{\mathrm{M}}$

and $T_{\mathrm{F}}$ can be represented as

$$T_{\mathrm{M}} = \sum_{r=1}^{N_{\mathrm{R}}} \sum_{m=1}^{N_{\mathrm{M}_r}} y_{\mathrm{M}_{rm}} \qquad (3)$$

$$T_{\mathrm{F}} = \sum_{r=1}^{N_{\mathrm{R}}} \sum_{m=1}^{N_{\mathrm{M}_r}} \sum_{f=1}^{N_{F_{mr}}} y_{\mathrm{F}_{rmf}}, \qquad (4)$$

where $y_{\mathrm{M}_{rm}}$ is the total cost incurred by municipality $m$ in region $r$ and $y_{\mathrm{F}_{rmf}}$ is the total cost specific to facility $f$ in municipality $m$ in region $r$. Note, this expression explicitly acknowledges that facilities are structured as belonging to some specific municipality within some specific region. Furthermore, the number of municipalities within any given region may vary across regions, and the number of facilities within any given municipality may also vary. Taking this structure into consideration is essential for valid estimation and inference.

## The sampling design and missing data

Generally, the costs incurred at region, municipality and facility levels are unknown. Researchers usually have access to readily available information on the hierarchical structure or demographic information at each level, but not to cost data. It is often the case that investigators have access or resources to collect detailed information on costs in a sub-sample of units. A cost-efficient approach is to implement a multi-stage sampling design where a sample of units is selected at each stage and the costs/variables of interest are recorded for each sampled unit. For example, in Brazil, the first stage units are regions and a sample of $n_{\mathrm{R}} = N_{\mathrm{R}}$ regions was selected. Because all regions were included in the sample, this is the same as a stratified sampling design. Subsequently, a sample of municipalities was selected from each region. The sampling design within each region was proportional to the number of children under one year old, with an expected sample size of $E(n_{\mathrm{M}_r})$, where $n_{\mathrm{M}_r}$ denotes the final number of municipalities sampled. Additionally, within each municipality, a random sample of $n_{\mathrm{F}_{rm}}$ facilities was selected. The expected municipality sample size per macro-region was 11 (55 in total, 5 macro regions), and the estimated facility sample size was 66 immunization services per macro-region, totaling 330 immunization services in the whole country. At each macro region, the facilities from the 11 selected municipalities were assembled into a pool. Then 66 facilities were selected from each pool. Due to this sampling mechanism, not all municipalities had facilities selected. In fact, this yielded a final sample of 330 vaccination units located in 40 Brazilian municipalities. Municipalities for which no health facilities were selected in the second stage of the sampling process were not visited and therefore did not have the municipal level data collected (15 municipalities). In addition, the information on two municipalities and its facilities was lost due to logistic issues. These municipalities were Inocência, in the central-west region, and Riachão do Jacupé in the Northeastern Region. This implies that the information from facilities belonging to these two municipalities was also missing.

Table 1(1) shows the final number of municipalities and facilities per region. For the Brazilian study, the total expected number of sampled municipalities was $E(n_{\mathrm{M}_r}) = 55$, but the actual size was $\sum_r n_{\mathrm{M}_r} = 38$, while the actual sample size of facilities was $\sum_{r,m} n_{\mathrm{F}_{rm}} = 325$.

## IPW (Inverse probability weighting) estimation

In absence of complete data at municipality and facility levels, neither $T_{\mathrm{R}}$, $T_{\mathrm{M}}$ nor $T_{\mathrm{F}}$ can be directly calculated. This is the case of multistage sampling design of our interest, where the

**Table 1. Final number of municipalities and facilities per region.**

| Region | No. mun. | No. mun. sampled | No. fac. | No. fac. sampled |
|---|---|---|---|---|
| Midwest | 464 | 10 | 1988 | 64 |
| Northeast | 1787 | 6 | 10897 | 64 |
| North | 447 | 9 | 2103 | 65 |
| Southeast | 1663 | 7 | 8312 | 66 |
| South | 1182 | 6 | 4087 | 66 |

total costs cannot be calculated. A possible approach for overcoming this limitation allowing for total costs estimation is the well-known inverse probability weighting (IPW) [10, 18]. It uses sampling weights to build a bridge between the observed subsample and the entire population to produce an estimator that represents all the population. Let $n_{F_{mr}}$ the total number of facilities sampled in municipality $m$ in region $r$ and $n_{F_r}$ be the total number of facilities sampled in region $r$. Focusing on $T_R$, the IPW estimate is given by

$$\hat{T}_R = \sum_{r=1}^{N_R} I_{R_r} w_r y_{R_r},$$

(5)

where $I_{R_r}$ is an indicator of whether the region $r$ was selected and $w_{R_r} = 1/\pi_{R_r}$, with $\pi_R$ the probability that region $r$ was selected to be part of the survey. In Brazil, for example, $I_{R_r} = 1$ for all five regions and $w_{R_r} = 1$. In some settings, such in the Honduras EPIC study, not all regions were included [3]. In such situations, some $I_{R_r}$ are zero. Similarly, an IPW estimator of $T_M$ is

$$\hat{T}_M = \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} I_{M_{rm}} w_{M_{rm}} y_{M_{rm}},$$

(6)

where $I_{M_{rm}}$ is an indicator of whether municipality $m$ in region $r$ was selected at the second stage sample, and $w_{M_{rm}} = 1/\pi_{M_{rm}}$ with $\pi_{M_{rm}}$ being the probability that the municipality was selected. In Brazil, 40 municipalities were selected initially, but the information on one of them is missing, then only 38 are available for the analysis.

Similarly, an IPW estimator of $T_F$ is

$$\hat{T}_F = \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} \sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} w_{F_{rmf}} y_{F_{rmf}},$$

(7)

where $I_{F_{rmf}}$ is an indicator of whether facility $f$ in municipality $m$, in region $r$ was selected at the third stage sample, and $w_{F_{rmf}} = 1/\pi_{F_{rmf}}$ with $\pi_{F_{rmf}}$ as the inclusion probability for such facility. The final IPW estimator is then given by $\hat{T} = T_C + \hat{T}_F + \hat{T}_M + \hat{T}_R$. Under a true design, statistical properties guarantee $\hat{T}$ to be an unbiased estimator of the total cost $T$ [10, 18].

## Variance estimation

Usually IPW estimators only involve data at the last stage of the sampling design. An example of this is $\hat{T}_F$. In this case, the variance of the estimator is well known and straight forward to calculate following standard properties of multi stage designs [18]. However, our estimator of interest involves information from the various stages of the sampling design. This implies that

the variance has to be estimated differently. Note first that

$$I_{F_{rmf}} = I_{F_{rmf}} I_{M_{rm}} I_{R_r}; \quad I_{M_{rm}} = I_{M_{rm}} I_{R_r}. \tag{8}$$

Using this, the number of facilities sampled in municipality $m$, region $r$ can be written as $n_{F_{rm}} = \sum_{f=1}^{N_{F_{rm}}} I_{F_{rmf}}$ and the number of facilities sampled in region $r$ can be written as $n_{F_r} = \sum_{m=1}^{N_{F_r}} I_{M_{rm}} \sum_{f=1}^{N_{F_{rm}}} I_{F_{rmf}} I_{F_{rmf}}$ and these quantities are fixed by design. Furthermore, note that $\hat{T}_R, \hat{T}_M, \hat{T}_F$ (Eqs (5), (6) and (7)) can respectively be written as

$$\hat{T}_R = \sum_{r=1}^{N_R} I_{R_r} \sum_{m=1}^{N_{M_r}} I_{M_{rm}} \sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} \frac{w_{F_{rmf}} w_{R_r} y_{R_r}}{w_{F_{rmf}} n_{F_r}}, \tag{9}$$

$$\hat{T}_M = \sum_{r=1}^{N_R} I_{R_r} \sum_{m=1}^{N_{M_r}} I_{M_{rm}} \sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} \frac{w_{F_{rmf}} w_{M_{rm}} y_{M_{rm}}}{w_{F_{rmf}} n_{F_{mr}}}, \tag{10}$$

$$\hat{T}_F = \sum_{r=1}^{N_R} I_{R_r} \sum_{m=1}^{N_{M_r}} I_{M_{rm}} \sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} w_{F_{rmf}} y_{F_{rmf}}. \tag{11}$$

Now, let

$$u_{rmf} = y_{F_{rmf}} + \frac{w_{M_{rm}}}{w_{F_{rmf}} n_{F_{mr}}} y_{M_{rm}} + \frac{w_{R_r}}{w_{F_{rmf}} n_{F_r}} y_{R_r}. \tag{12}$$

Then, using (9)–(11) and $u_{rmf}$, we can write the estimator $\hat{T}$ as

$$\hat{T} = \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} \sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} w_{F_{rmf}} u_{rmf}. \tag{13}$$

Since $n_{F_r}$, $n_{F_{rm}}$, $w_{M_{rm}}$ and $w_{F_{rmf}}$ are fixed by design, all the design-based uncertainty is captured in the expression above and we can calculate the variance of the estimator using standard expressions or techniques for the IPW estimators, but now the target estimator is for the total of the variable $u_{rmf}$. Let $s_R$, $s_{M_r}$ and $s_{F_{rm}}$ denote the sample of regions, municipalities in region $r$ and facilities in municipality $m$ in region $r$, respectively. Note that the variance of the estimator can be calculated using the expression

$$Var(\hat{T}) = E[Var(\hat{T}|u)] + Var[E(\hat{T}|u)]. \tag{14}$$

Following [18], the expression for the first term is

$$E[Var(\hat{T}|u)] = E\left[\sum_{r=1}^{N_R} \sum_{s=1}^{N_R} \Delta_{R_{rs}} \check{t}_{R_s} \check{t}_{R_r} + \sum_{r=1}^{N_R} w_{R_r} V_{R_r}\right], \tag{15}$$

where $\Delta_{R_{rs}} = \pi_{R_{rs}} - \pi_{R_r} \pi_{R_s}$, $\pi_{R_{rs}} = P(I_{R_r} = 1, I_{R_s} = 1)$ which represents the pairwise inclusion probabilities at the regional level, $\check{t}_{R_r} = w_{R_r} \sum_{m=1}^{N_{F_r}} \sum_{f=1}^{N_{F_{rm}}} u_{rmf}$. The term $V_{R_r}$ is given by $V_{R_r} = Var(\hat{t}_{R_r}|u)$, where $\hat{t}_{R_r} = \sum_{m=1}^{N_{M_r}} \sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} w_{F_{rmf}} \pi_{R_r} u_{rmf}$. The term $V_{R_r}$ is calculated in the same way as (15) and so on for each sampling stage. Further details are given in the supporting

information S1. An estimator of (15) is given by

$$\hat{E}\big[Var(\hat{T}|u)\big] = \sum_{r=1}^{N_R}\sum_{s=1}^{N_R} I_{R_s} I_{R_r} \frac{\Delta_{R_{rs}}}{\pi_{R_{rs}}} \frac{\hat{t}_{R_s}}{\pi_{R_s}} \frac{\hat{t}_{R_r}}{\pi_{R_r}} + \sum_{r=1}^{N_R} I_{R_r} w_{R_r} \hat{V}_{R_r}, \tag{16}$$

where $\hat{V}_{R_r}$ is such that $E[\hat{V}_{R_r}|u, s_R] = V_{R_r}$. This expression is calculated in a similar way to (16). Details on this derivations can be found in [18]. We also provide specific formulas for (15) and (16) for the Brazilian design. And, an unbiased estimator for the second term in 14 is given by

$$\widehat{Var}\big[E(\hat{T}|u)\big] = \sum_{r=1}^{N_R} I_{R_r} \sum_{m=1}^{N_{M_r}} I_{M_{rm}} \frac{N_R N_{M_r}}{n_R n_{M_r}} N_{F_{mr}} \hat{S}^2_{u_{|mr}}, \tag{17}$$

where

$$\hat{S}^2_{u_{|mr}} = \frac{1}{n_{F_{mr}}-1}\sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} (u_{rmf} - \tilde{u}_{|m_r})^2; \quad \tilde{u}_{|m_r} = \frac{1}{n_{F_{mr}}}\sum_{f=1}^{N_{F_{mr}}} I_{F_{rmf}} u_{rmf}. \tag{18}$$

## Estimation of the cost per dose

The cost per dose is defined as the total cost of implementation divided by the number of doses provided by the program. Mathematically, $c_p = \frac{T}{T_d}$, where $T$ is the total cost of the program and $T_d = \sum_{r=1}^{N_R}\sum_{m=1}^{N_{M_r}}\sum_{f=1}^{N_{F_{mr}}} d_{rmf}$ is the total number of doses, with $d_{rmf}$ being the number of doses in facility $f$ in municipality $m$ in region $r$. The estimate of the cost per dose is

$$\hat{c}_p = \frac{\hat{T}}{T_d}, \tag{19}$$

where $\hat{T}$ is the estimated total cost obtained in (13). The variance of this estimator is given by

$$Var(\hat{c}_p) = \frac{Var(\hat{T})}{T_d^2} \tag{20}$$

and an unbiased estimator of this quantity is given by

$$\hat{Var}\left(\hat{c}_p\right) = \frac{\hat{V}(\hat{T})}{T_d^2}, \tag{21}$$

where $\hat{V}(\hat{T})$ is given by (14).

## Post-design methods: Calibrated weights

The main purpose of post-designs methods is to improve the estimates by reducing the variance, i. e. to produce more efficient estimators. Post-design methods include techniques such as post stratification, estimation of weights and calibration of weights. The latter is known to yield more efficient estimates by adjusting the weights using information readily-available at first phase. In Brazil, for example, the total number of doses administered and the total number of facilities in the country is known. Hence this can be used to adjust (calibrate) the weights.

[13] proposed the use of calibration to adjust the sampling weights in case–cohort designs. Let $v_F$, $v_M$ and $v_R$ denote readily available variables at first phase and at the different levels. The former is a variable available for all units at the final stage, the second is available for all second

stage units and the latter is available for all first-stage units, e.g regions. The main goal is to adjust weights such that estimates for the totals $\hat{T}_{v_F}$, $\hat{T}_{v_M}$ and $\hat{T}_{v_R}$ equal the actual totals $T_{v_F}$, $T_{v_M}$ and $T_{v_R}$, but forcing the new weights ($\hat{w}$) to be as close as possible to the design weights. For example, consider IPW estimates $\hat{T}_{v_F}$ and $\hat{T}_{v_M}$ based on the same weights used to estimate $T$ with $\hat{T}$, i.e.

$$\hat{T}_{v_F} = \sum_{s_{F_{rm}}} w_{F_{rmf}} v_{Frmf}; \quad \hat{T}_{v_M} = \sum_{s_{M_r}} w_{M_{rm}} v_{Mrm}; \quad \hat{T}_{v_R} = \sum_{s_R} w_{R_r} v_{Rr}.$$

Since we know $T_{v_F}$, $T_{v_M}$ and $T_{v_R}$ we can slightly modify the weights ($w$) and find ($\hat{w}$) such that

$$
\begin{aligned}
\hat{T}_{v_F} &= \sum_{s_{F_{rm}}} \hat{w}_{F_{rmf}} v_{Frmf} = T_{v_F}; \\
\hat{T}_{v_M} &= \sum_{s_{M_r}} \hat{w}_{M_{rm}} v_{Mrm} = T_{v_M}; \\
\hat{T}_{v_R} &= \sum_{s_R} \hat{w}_{R_r} v_{Rr} = T_{v_R}.
\end{aligned}
\tag{22}
$$

Intuitively, the modified weights ($\hat{w}_{F_{rmf}}$, $\hat{w}_{M_{rm}}$) contain information about the entire population of facilities and we can use the new weights to find a new estimate of $T$:

$$\hat{T}_{cal} = \hat{T}_{F,cal} + \hat{T}_{M,cal} + \hat{T}_{R,cal} + T_C = \sum_{s_{F_{rm}}} \hat{w}_{F_{rmf}} y_{F_{rmf}} + \sum_{s_{M_r}} \hat{w}_{M_{rm}} y_{M_{rm}} + \sum_{s_R} \hat{w}_{R_r} y_{R_r} + T_C,$$

where $\hat{T}_{F,cal}$, $\hat{T}_{M,cal}$ and $\hat{T}_{R,cal}$ represent the estimators of total cost at the Facility, Municipality and regional level, respectively. Mathematically, a distance function $d(w, \hat{w})$ measuring the distance from original weights($w$) to new weights($\hat{w}$) is minimized subject to the constrains (22). There are several options for choosing this distance. Substantial gains in accuracy depend on how correlated the calibration variables are with the quantity of interest [9].

## Variance estimation

In order to estimate the variance in practice we use a Taylor expansion as in [18]. Define

$$\tilde{u}_{rmf} = e_{F_{rmf}} + \frac{w_{M_{rm}}}{w_{F_{rmf}} n_{F_{mr}}} e_{M_{rm}} + \frac{w_{R_r}}{w_{F_{rmf}} n_{F_r}} e_{R_r},$$

where $e_{F_{rmf}}$, $e_{M_{rm}}$ and $e_{R_r}$ y are given by

$$
\begin{aligned}
e_{F_{rmf}} &= y_{F_{rmf}} - v_{Frmf}^T \tilde{\boldsymbol{B}}_F; \quad \tilde{\boldsymbol{B}}_F = \tilde{\boldsymbol{T}}_F^{-1} \left( \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} \sum_{f=1}^{N_{F_{m_r}}} v_{R_{rmf}} \right) \\
e_{M_{rm}} &= y_{M_{rm}} - v_{Mrm}^T \tilde{\boldsymbol{B}}_M; \quad \tilde{\boldsymbol{B}}_M = \tilde{\boldsymbol{T}}_M^{-1} \left( \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} v_{M_{rm}} \right) \\
e_{R_r} &= y_{R_r} - v_{Rr}^T \tilde{\boldsymbol{B}}_R; \quad \tilde{\boldsymbol{B}}_R = \tilde{\boldsymbol{T}}_R^{-1} \left( \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} v_{R_{rm}} \right),
\end{aligned}
\tag{23}
$$

**Table 2. Information used for calibration of the weights.** FAcility size is defined as follows. Huge (No.Doses > 10000); large (5000<No.Doses ≤ 10000); medium (1500<No.Doses ≤ 5000); small (500 < No.Doses ≤ 1500) and tiny (No.Doses ≤ 500).

|  | Huge | Large | Medium | Small | Tiny | Total Doses |
|---|---|---|---|---|---|---|
| Frame | 1727 | 3210 | 10917 | 8260 | 3273 | 89928432 |
| Sample | 77 | 61 | 106 | 47 | 34 | 2822574 |
|  | Midwest | Northeast | North | Southeast | South | No. Municipalities |
| Frame | 1988 | 10897 | 2103 | 8312 | 4087 | 5543 |
| Sample | 64 | 64 | 65 | 66 | 66 | 38 |

https://doi.org/10.1371/journal.pone.0212401.t002

and

$$\tilde{w}_{F_{rmf}} = w_{F_{rmf}}(1 + v_{F_{rmf}}^{\mathrm{T}}\tilde{\lambda}_{\mathrm{R}}); \quad \tilde{\lambda}_{\mathrm{R}} = \tilde{T}_{\mathrm{R}}^{-1}(T_{v_{\mathrm{F}}} - \hat{T}_{v_{\mathrm{F}}}); \quad \tilde{T}_{\mathrm{F}} = \sum_{r=1}^{n_{\mathrm{R}}}\sum_{m=1}^{n_{\mathrm{M}_r}}\sum_{f=1}^{n_{F_{m_r}}}\tilde{w}_{F_{rmf}}v_{F_{rmf}}v_{F_{rmf}}^{\mathrm{T}};$$

$$\tilde{w}_{M_{rm}} = w_{F_{rm}}(1 + v_{M_{rm}}^{\mathrm{T}}\tilde{\lambda}_{\mathrm{M}}); \quad \tilde{\lambda}_{\mathrm{M}} = \tilde{T}_{\mathrm{M}}^{-1}(T_{v_{\mathrm{M}}} - \hat{T}_{v_{\mathrm{M}}}); \quad \tilde{T}_{\mathrm{M}} = \sum_{r=1}^{n_{\mathrm{R}}}\sum_{m=1}^{n_{\mathrm{M}_r}}\tilde{w}_{M_{rm}}v_{M_{rm}}v_{M_{rm}}^{\mathrm{T}}; \quad (24)$$

$$\tilde{w}_{R_r} = w_{R_r}(1 + v_{R_r}^{\mathrm{T}}\tilde{\lambda}_{\mathrm{R}}); \quad \tilde{\lambda}_{\mathrm{F}} = \tilde{T}_{\mathrm{F}}^{-1}(T_{v_{\mathrm{F}}} - \hat{T}_{v_{\mathrm{F}}}); \quad \tilde{T}_{\mathrm{R}} = \sum_{r=1}^{n_{\mathrm{R}}}\tilde{w}_{R_r}v_{R_r}v_{R_r}^{\mathrm{T}}.$$

Therefore

$$\widehat{Var}(\hat{T}_{cal}) = Var(\hat{T}_e) = \hat{E}[Var(\hat{T}_e|u)] + \widehat{Var}[E(\hat{T}_e|u)], \quad (25)$$

with $\hat{T}_e = \sum_{r=1}^{N_{\mathrm{R}}}\sum_{m=1}^{N_{\mathrm{M}_r}}\sum_{f=1}^{N_{F_{mr}}} I_{R_r} I_{M_{rm}} I_{F_{rmf}} w_{F_{rmf}} \tilde{u}_{rmf}$. Expression (25) is found in a similar way as the variance estimate in section Variance estimation.

## Results from Brazilian study

Table 2 displays the information used for calibration of the weights. At the facility level, weights ($w_F$) were calibrated using all the totals the totals presented in Table 2. This includes information on the facility size, the total number of doses in 2014 and the number of facilities in each region. Furthermore, at the municipality level, weights($w_M$) were calibrated using only the total number of municipalities. The number of expected municipalities in the sample was 55, but the actual sample contained 38 municipalities. This was due to sampling variation and due to two missing municipalities.

Tables 3 and 4 display the estimated total cost for each of the different categories of cost at the municipal level and for the entire program, respectively. Table 3 presents the estimates of costs incurred by municipalities. The data used to estimate this was the sample of 38 munici-palities. The table shows significant differences between the estimates using the unadjusted sampling weights and the estimates using the calibrated weights. The reason for this is the small sample of municipalities available. With this in mind, adjusting/ calibrating the sampling weights yields more representative and informative estimators. In addition, due to the increase in magnitude, the SE are larger for the calibrated estimators.

Table 4 displays the estimates of the total cost of the entire program for each of the different categories. As observed, the estimates differ in magnitude for all the categories and the cali-brated estimates are always larger. An additional feature of the calibrated estimates is that their standard errors are lower than those obtained with the unadjusted sampling weights. This

leads to the conclusion that the calibrated estimates are more precise and efficient and therefore more reliable.

## Results

**Table 3. Estimated municipality level costs, by cost category, considering unadjusted and calibrated weight (in million R$) for the Costing Study of the Brazilian Immunization Program.** Brazil, 2013. Standard errors are displayed in parenthesis. Municipal level weights were calibrated to the national number of municipalities.

| Category | Sampling weights | | | Calibrated weights | | |
|---|---|---|---|---|---|---|
| | Capital | Recurrent | Total | Capital | Recurrent | Total |
| Vehicles | 43(14) | 4.3(1.4) | 47(15.1) | 202 (71) | 20.2(7.1) | 222(78.1) |
| Equipment | 3.8(0.6) | 0.4(0.1) | 4.2(0.7) | 18.1(4.7) | 1.8(0.5) | 20(5.2) |
| Buildings | 9.9(1.7) | 72(44.7) | 82.1(45) | 47(12.9) | 341(233) | 387(239) |
| Labor | | 125(21.6) | 126(22) | | 593(123) | 593(123) |
| Other | | 12(5.7) | 12(5.7) | | 56.5(29.5) | 56.5(29.5) |
| Total | 57(14.8) | 215(51.5) | 271(54) | 267(82) | 1012(314) | 1279(355) |

## Discussion

We have proposed inverse probability weighting(IPW) with calibration methods to improve estimators in studies where information is collected at several levels, e.g. states-municipalities-clinics. We find an expression for the estimator of the total such that it accounts for the totals at each level and also provides a feasible way to find the variance. Further, we use information available for all units in the targeted population to adjust the weights with calibration methods. It allows for calibration of the weights at all the levels of sampling if desired by investigators. These methods can be used to correct for missing information at the various levels. This is achieved by adjusting the estimates at each level to comply with the characteristics of the country. The methods were applied to the Brazilian National Immunization Program. The results showed that calibrating the weights does not only correct for missingness, but also results in large gains in efficiency and precision compared to standard methods. The estimators result difficult to compare in terms of efficiency due to the difference in magnitude. To provide an interpretation, we considered that the estimates obtained by calibration were consistent, which is supported by asymptotic theory [9, 10]. We compare the estimated MSEs as follows. For the total cost with standard weights, the MSEs are: 278(Capital), 1138(Recurrent), 1415(Total). With calibrated weights these are 61(Capital), 291 (Recurrent), 331(Total). Calibration clearly returns smallest MSEs if the estimator is consistent (or unbiased). A point of interest for future research is to investigate how different would the results have been if the sampling design had

**Table 4. Estimated total costs, by cost category, considering unadjusted and calibrated weight (in million R$) for the Costing Study of the Brazilian Immunization Program.** Brazil, 2013. Standard errors are displayed in parenthesis. This includes facility level cost, municipality level cost and state level cost. Municipal level weights were calibrated to the national number of municipalities, and facility level weights were calibrated using information presented in Table 2.

| Category | Sampling weights | | | Calibrated weights | | |
|---|---|---|---|---|---|---|
| | Capital | Recurrent | Total | Capital | Recurrent | Total |
| Vehicles | 45.1 (6.8) | 8.5 (1.1) | 53.5 (7.8) | 248 (50.3) | 27.8 (5.2) | 275.8 (55.5) |
| Equipment | 25.6 (3.4) | 4.2 (0.6) | 29.8 (4) | 46.3 (5.9) | 6.3 (0.8) | 52.6 (6.6) |
| Buildings | 83.1 (8.9) | 94.7 (32.3) | 177.8 (35.2) | 137.5 (12.2) | 466 (165.8) | 604 (169.3) |
| Labor | | 1310 (122.2) | 1310 (122.2) | | 1956 (186.2) | 1956(186.2) |
| Other | | 16.9 (3.4) | 16.9 (3.4) | | 107.1 (34.3) | 107.1 (34.3) |
| Total | 153.8 (16.2) | 1434 (136.3) | 1588 (149.8) | 431.8 (60.8) | 2564 (291.4) | 2996 (331.4) |

been different. For example, stratifying by facility size and/or type. This can enhance estimators because facility size/type is related to the cost. On the contrast, stratification by something not related to the cost would not yield gains in efficiency. We hope that investigators become more aware of the existence and implementation of these methods. This can help analyst and policy-makers to inform on the basis of more reliable information.

## Supporting information

**S1 Appendix. Estimation and variance of the total cost and cost per dose from the Brazilian immunization study.**
(PDF)

**S2 Appendix. Estimation and variance of the total cost and cost per dose using design and calibrated weights.**
(PDF)

**S1 Dataset. The dataset contains the data for the estimation of the total cost from the Brazilian study.**
(TXT)

## Author Contributions

**Conceptualization:** Claudia Rivera-Rodriguez.

**Formal analysis:** Claudia Rivera-Rodriguez.

**Methodology:** Claudia Rivera-Rodriguez.

**Project administration:** Cristiana Toscano.

**Supervision:** Cristiana Toscano, Stephen Resch.

**Writing – original draft:** Claudia Rivera-Rodriguez.

**Writing – review & editing:** Cristiana Toscano.

## References

1. Brenzel L, Schütte C, Goguadze K, Valdez W, Le Gargasson JB, Guthrie T. EPIC Studies: Governments Finance, On Average, More Than 50 Percent Of Immunization Expenses, 2010-11. Health Affairs. 2016; 35(2):259–265. https://doi.org/10.1377/hlthaff.2015.1121 PMID: 26858378

2. Brenzel L, Young D, Walker DG. Costs and financing of routine immunization: Approach and selected findings of a multi-country study (EPIC). Vaccine. 2015; 33 Suppl 1:A13–20. https://doi.org/10.1016/j.vaccine.2014.12.066 PMID: 25919153

3. Janusz CB, Castañeda-Orjuela C, Molina Aguilera IB, Felix Garcia AG, Mendoza L, Díaz IY, et al. Examining the cost of delivering routine immunization in Honduras. Vaccine. 2015; 33 Suppl 1:A53–9. https://doi.org/10.1016/j.vaccine.2015.01.016 PMID: 25919175

4. Robson M, Andrus JK, Toscano CM, Lewis M, Oliveiria L, Ropero AM, et al. A Model for Enhancing Evidence-Based Capacity to Make Informed Policy Decisions on the Introduction of New Vaccines in the Americas: Paho's Provac Initiative. Public Health Reports. 2007; 122(6):811–816. https://doi.org/10.1177/003335490712200613

5. Le Gargasson JB, Nyonator FK, Adibo M, Gessner BD, Colombini A. Costs of routine immunization and the introduction of new and underutilized vaccines in Ghana. Vaccine. 2015; 33 Suppl 1:A40–6. https://doi.org/10.1016/j.vaccine.2014.12.081 PMID: 25919173

6. Fundo Nacional de Saude F. Study Report: Brasil. Estimativa de custo do programa nacional de imunizacao. Relatorio Final. Ministerio da Saude, Programa Nacional de Imunizacao; 2018.

7. Schütte C, Chansa C, Marinda E, Guthrie TA, Banda S, Nombewu Z, et al. Cost analysis of routine immunisation in Zambia. Vaccine. 2015; 33 Suppl 1:A47–52. https://doi.org/10.1016/j.vaccine.2014.12.040 PMID: 25919174

8. Goguadze K, Chikovani I, Gaberi C, Maceira D, Uchaneishvili M, Chkhaidze N, et al. Costs of routine immunization services in Moldova: Findings of a facility-based costing study. Vaccine. 2015; 33 Suppl 1:A60–5. https://doi.org/10.1016/j.vaccine.2014.12.034 PMID: 25919177

9. Deville JC, Sarndal CE. Calibration estimators in survey sampling. Journal of the American Statistical Association. 1992; 87:376–382. https://doi.org/10.1080/01621459.1992.10475217

10. Fuller W. Sampling Statistics. John Wiley and Sons; 2009.

11. Lumley T, Shaw PA, Dai JY, Tsiatis AA, Davidian M, Handcock MS, et al. Connections between Survey Calibration Estimators and Semiparametric Models for Incomplete Data [with Discussions]. International Statistical Review / Revue Internationale de Statistique. 2011; 79(2):200–232. https://doi.org/10.1111/j.1751-5823.2011.00138.x PMID: 23833390

12. Breslow N and Amorim G and Pettinger M and Rossouw J. Using the Whole Cohort in the Analysis of Case-Control Data: Application to the Women's Health Initiative. Stat Biosci. 2014; 5(2). https://doi.org/10.1007/s12561-013-9080-2

13. Breslow NE, Lumley TS, Ballantyne CM, Chambless LE, Kulich M. Using the whole cohort in the analysis of case-cohort data. American journal of epidemiology. 2009; 169(11):1398–405. https://doi.org/10.1093/aje/kwp055 PMID: 19357328

14. Breslow N, Lumley T, Ballantyne C, Chambless L, Kulich M. Improved Horvitz-Thompson Estimation of Model Parameters from Two-phase Stratified Samples: Applications in Epidemiology. Statistics in Biosciences. 2009; 1(1):32–49. https://doi.org/10.1007/s12561-009-9001-6 PMID: 20174455

15. Haneuse S, Hedt-Gauthier B, Chimbwandira F, Makombe S, Tenthani L, Jahn A. Strategies for monitoring and evaluation of resource-limited national antiretroviral therapy programs: the two-phase design. BMC Medical REsearch Methodology. 2015; 15(31). https://doi.org/10.1186/s12874-015-0027-9 PMID: 25886976

16. Kim JK, Kwon Y, Paik MC. Calibrated propensity score method for survey nonresponse in cluster sampling. Biometrika. 2016; 103(2):461. https://doi.org/10.1093/biomet/asw004 PMID: 27279670

17. Walker DG, Hutubessy R, Beutels P. WHO Guide for standardisation of economic evaluations of immunization programmes. Vaccine. 2010; 28(11):2356–2359. https://doi.org/10.1016/j.vaccine.2009.06.035 PMID: 19567247

18. Sarndal CE, Swensson B, Wretman J. Model Assisted Survey Sampling. New York; 1992.