Prevalence and number of children living in institutional care: global, regional, and country estimates

Chris Desmond, Kathryn Watt, Anamika Saha, Jialin Huang, Chunling Lu

Lancet Child Adolesc Health 2020; 4: 370–77

Published Online March 6, 2020 https://doi.org/10.1016/S2352-4642(20)30022-5

See Comment page 343

Centre for Liberation Studies, Durban, KwaZulu-Natal, South Africa (C Desmond PhD, K Watt MA); **Centre for Rural Health, University of KwaZulu-Natal, Durban, KwaZulu-Natal, South Africa** (C Desmond); **Harvard School of Public Health, Boston, MA, USA** (A Saha MPH); **National School of Development, Peking University, Beijing, China** (J Huang BA); **and Division of Global Health Equity, Brigham and Women's Hospital, and Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA** (C Lu PhD)

Correspondence to: Dr Chunling Lu, Brigham & Women's Hospital—Harvard Medical School, Boston, MA 02115, USA
chunling_lu@hms.harvard.edu

Summary

Background Children living in institutionalised settings are at risk of negative health and developmental outcomes, as well as physical and emotional abuse, yet information on their numbers is scarce. Therefore, the aim of our study was to estimate global-level, regional-level, and country-level numbers and percentages of children living in institutional care.

Methods In this estimation study, we did a systematic review of peer-reviewed publications and a comprehensive review of surveys and unpublished literature to construct a dataset on children living in institutional care from 136 countries between 2001 and 2018. We applied a wide range of methods to estimate the number and percentages of children living in institutional care in 191 countries in 2015, the year the Sustainable Development Goals were adopted. We generated 98 sets of estimates for each dataset, with possible combinations of imputation methods for countries with different available data points. Of these 98 sets, we report here five types of global-level estimates: estimates with the highest values, those with the lowest values, those with median values, those with uncertainty levels, and those derived from methods with the smallest root-mean-square errors (RMSE).

Findings Global estimates of children living in institutions in 2015 was highly sensitive to the methods and data used, ranging from 3·18 million to 9·42 million children, with a median estimate of 5·37 million. When selecting the method with the lowest RMSE, the global estimate was 4·21 million, whereas with negative binomial regression with bootstrapping, the global estimate was 7·52 (95% CI 7·48–7·56) million. We also observed large variations in country-level estimates. Compared with other regions, estimates in south Asia, sub-Saharan Africa, and Latin America had larger variations in values when switching between estimation methods. High-income countries had the highest average prevalence of institutionalisation, whereas low-income countries had the lowest average prevalence. Estimates from the full data with the smallest RMSE method showed that south Asia had the largest estimated number of children living in institutions (1·13 million), followed by Europe and central Asia (1·01 million), east Asia and Pacific (0·78 million), sub-Saharan Africa (0·65 million), Middle East and North Africa (0·30 million), Latin America and the Caribbean (0·23 million), and North America (0·09 million). North America consistently had the lowest estimates among all regions.

Interpretation Worldwide, institutional care places millions of children at elevated risk of negative health and developmental outcomes, highlighting the need for deinstitutionalisation. However, there is considerable uncertainty regarding the number of children living in institutions. To improve estimates of the size of this population, we need to standardise the definition of institutional care and improve data collection, particularly in countries with large child populations.

Funding Lumos Foundation.

Introduction

Children living in institutionalised settings (eg, orphanages, residential homes) receive impersonal care that is often of poor quality, are easily exploited, are at risk of physical and emotional abuse, and might have unmet medical needs.[1–6] The nature of such care, particularly when combined with other adverse childhood experiences, can impair children's health (eg, growth suppression and dysregulation of the hypothalamic-pituitary-adrenal system) and development (eg, neurological, cognitive, language, and social development).[7–14] Potential long-term consequences include reduced intelligence quotient and elevated risk of metabolic conditions in adulthood, such as hypertension and heart disease.[1,14]

Monitoring the number of children living in these high-risk contexts is key to assessing progress towards de-institutionalisation. However, not all countries record the number of children living in institutions and, among those that do, access to data or quality of data are often a concern.[15]

The most recent study estimated that approximately 2·7 million children (from birth to 17 years of age) lived in institutional care settings worldwide in 2012.[16] The estimate was based only on official government-reported data, and no country-level data or estimates were published for review. The study recognised the challenges of capturing and recording the number of children in institutional care and noted the risks of

Research in context

Evidence before this study

We systematically searched PubMed for peer-reviewed studies published from Jan 1, 2008, to Sept 30, 2018, on global-level, regional-level, and country-level estimates of children living in institutions. A list of the search terms used is available in the appendix (p 4). Additionally, we did a comprehensive search of publicly available datasets and data from unpublished literature to supplement the review of published studies. We found 130 peer-reviewed publications, most of which cited official data or were counts of subnational, regional, or global estimates with no information on the methods provided. Only one study provided a global estimate with details on methods and data used; this study used official data, which is often thought to underestimate prevalence, and did not make country-level estimates available.

Added value of this study

This study makes two major contributions to the existing knowledge about children living in institutions. First, through a comprehensive search of literature and information obtained from various sources, we compiled a unique dataset of 344 data points from 136 countries during the period of 2001–18 and provided a comprehensive assessment of the data quality. Second, using the compiled country-level data, we did an estimation at the global, regional, and country level and presented how using different estimation methods, data from diffusing sources, or data with different definitions could affect estimation by generating 98 sets of estimates for each dataset. With this study, we provide a resource that could help facilitate informed debate on the size of this vulnerable population.

Implications of all the available evidence

These findings highlight that, despite the known risks to health and development, millions of children globally are placed in institutional care, suggesting that existing efforts towards deinstitutionalisation are inadequate. The uncertainty of the estimates highlights the clear need for agreement on a standard definition of institutional care and more investment in data collection and validation efforts to improve the availability, reliability, and comparability of official data on children living in institutional care. Better evidence on the number of children living in institutional care will help to drive targeted country-level efforts towards deinstitutionalisation.

See Online for appendix

relying on official data, particularly in low-income and middle-income countries where information systems are generally weak.

Our study aimed to improve the estimation of the number of children living in institutions at the global, regional, and country level by constructing a comprehensive dataset comprising information obtained from official and non-official sources and applying a wider range of statistical techniques than that used to make the 2015 estimate. Additionally, we aimed to investigate the sensitivity of estimates to different methods or datasets used.

Methods

Definition of children living in institutional care

No standard definition of institutional care exists.[17] Even the UN guidelines on alternative care do not offer one, beyond equating institutions with "large residential care facilities", with no indication of what constitutes a large facility.[18] Efforts to define institutional care have focused on the size of establishments (number of children cared for) or the nature of care (eg, the use of paid shift staff and structured programmes).[17] Which characteristics are considered most salient determines whether a particular alternative care setting is defined as an institution. For example, when focusing on the number of children, small group homes would not be considered institutional care, whereas focusing on paid caregiving would classify them as such.

When compiling our dataset, we included estimates from settings referred to as institutions, institutional care, public or societal care, residential care, orphanages, and children's homes. We excluded estimates from settings referred to as foster care. We noted substantial variation in the use of the term residential care. At times, this term included specialised boarding schools, which act as social care institutions, and long-term health facilities, whereas at other times, these were excluded. By including residential care in this way, we relied on the judgment of the people compiling the statistics: for example, if they reported long-term health facilities in their context alongside orphanages, it is because they considered them a similar form of care. We included family-like care (eg, small group homes) when it was included with estimates of other forms of institutional care. To the extent possible, we removed estimates of the number of children in the criminal justice system. We included estimates of children in both formal (state or court placement) and informal (non-state placement) care arrangements. Additional details on which definitions were included are provided in the appendix (pp 1–3).

Data sources

We compiled a dataset on the number of children living in institutions by extracting data from multiple sources: a systematic search of peer-reviewed publications, a comprehensive search of the grey literature (eg, reports produced by non-governmental organisations or government documents), a review of surveys in developing countries hosted by the International Household Survey Network,[19] and a 2018 survey done by The Global Study of

No data (55 countries)	One data point (54 countries)	Two or more data points (82 countries)
• Regression prediction • Average by income group across all years • Multiple regression	• Regression prediction • Average by income group across all years • Multiple regression • Using available data	• Regression prediction • Average by income group across all years • Linear extrapolation • Closest year • Average by country • Lowest by country • Highest by country • Multiple regression

Figure 1: **Imputation strategies**

Children Deprived of Liberty (GSCDL). Details of the data gathering process and summary statistics of the collected data are shown in the appendix (pp 4–14).

This compilation yielded 344 data points in 136 countries during the period of 2001–18. The national-level data on children in institutions were obtained from peer-reviewed literature (four countries), grey literature (129 countries), the Integrated Public Use Microdata Series (census data in ten countries that included information on institutions or group quarters), and the GSCDL survey (24 countries; appendix p 10). Because data were obtained from different sources, 24 countries had multiple data points in 1 year. We selected the lowest and highest estimates in the year for those 24 countries and split the data into two sets, one with the lower values and the other with the higher values. We used the dataset with the higher values in our estimation and report the results here. The dataset with the lower values was used for sensitivity tests (appendix pp 29–52).

Data were identified for 37 high-income countries, 39 upper-middle-income countries, 31 lower-middle-income countries, and 29 low-income countries, as classified by the World Bank in 2018 (appendix p 10).[20] The median year of the estimates was 2012 (IQR 2007–2015).

For any given year, the number of countries with data available was small. 2017 had the highest coverage in all years during the period, but children in countries with data accounted for only 29% of all children (appendix p 12). A summary of the data collected by region and by income group is provided in the appendix (pp 12–13).

Ethical approval was not required because the data are anonymous and in the public domain.

Estimating prevalence and total number of children living in institutions

To generate global-level, regional-level, and country-level estimates of the prevalence of children living in institutions, we selected 191 countries with available data on income classification, region, and percentage of population younger than 18 years. We chose 2015 as the estimation year because it was one of the three years (2007, 2014, and 2015) in which many countries (35 in 2007 and 2014 and 34 in 2015) had available data on children living in institutions. Of the three years, 2015 had the highest coverage of available data (21·4%, appendix p 12). Moreover, 2015 was the final year of the Millennium Development Goal era; therefore, if comparable data are collected in the future, estimates in 2015 could be used as a baseline for monitoring the progress in the Sustainable Development Goal era.

We divided the 191 countries into three categories: 55 countries with no data on children living in institutions between 2001 and 2018 (10·9% of the total population of children in 2015), 54 countries with one data point (24·7% of the total population of children in 2015), and 82 countries with two or more data points (64·4% of the total population of children in 2015).

For countries in these categories, we explored various strategies to impute missing values in 2015 (figure 1). For countries with no data during the period, we used three methods: predicting the missing values by use of regression models, assigning the average percentage of children living in institutions during the period by income group to countries in the corresponding income group, and doing multiple imputation. For the regression analysis, we used a generalised linear model with a log link and binomial distribution; this model is recommended when the dependent variable comprises data reported in proportions.[21] We also used models for count data by including child population as an offset variable.[22] A negative binomial model was chosen because it allows for the relaxation of Poisson assumption that the mean and variance are the same.[23] We used a bootstrapping method to generate the mean (95% CI) for the estimates derived from regression analysis.[24] To process the regression prediction and multiple imputation, we obtained a group of predictors from the World Bank and selected those with a relatively high correlation with the percentage of children in institutions. Details on the pattern of missing data, predictor selection, and implementation of imputation strategies are provided in the appendix (pp 15–26).

For countries with only one data point during 2001–18, but not in 2015, we used these described methods. Additionally, we used the available data in a country during the period to impute its missing data in 2015.

For countries with no data in 2015, but with two or more data points during 2001–18, beyond the described methods, we used five additional methods to replace the missing data: linear extrapolation, using the existing data in the closest year, and taking the average, lowest, or highest value of the existing data in each country during the period (figure 1).

To investigate the sensitivity of the estimation, we applied these imputation strategies to the full data (136 countries with available data, higher *vs* lower values), and to data obtained from or attributed to official sources only (103 countries with available data, higher *vs* lower values). We also applied the methods from the 2016 study by UNICEF on estimating the global number of children

living in institutions, meaning that we used the available official data in the most recent years to obtain weighted regional averages of prevalence of children in institutions, and applied them to the number of children in each region to obtain global-level estimates.[16]

We generated country-level estimates of number of children living in institutions in 2015 by multiplying a country's existing or imputed percentage by its number of children (data sources are given in the appendix, p 10). We also derived estimates of prevalence and number of children living in institutions at the regional and global level, as well as by income group.

For each dataset, we generated 98 sets of estimates with possible combinations of imputation methods for countries with different available data points. Details on how we used those methods are presented in the appendix (pp 15–26). Of those 98 sets, we report in the text five types of global-level estimates: those with the highest values, those with the lowest values, those with median values, those with uncertainty levels, and those derived from methods with the smallest root-mean-square errors (RMSE).[25]

RMSE has been applied in similar studies to select between estimates made with different methods.[26] It measures the difference between predicted values and observed values and is calculated as the square root of the mean of the squares of the deviations between predicted values and observed values. The smaller the RMSE is, the closer the imputed values are to the observed values. Assuming data are missing at random (appendix p 19), for countries without any data on children living in institutions, predicted values from negative binomial regressions yielded the smallest RMSE; for countries with only one data point, using available data yielded the smallest RMSE; for countries with two or more data points during the period, using the values from linear extrapolation yielded the smallest RMSE. Details on calculating the RMSE are presented in the appendix (pp 27–28). Because of the high rate of missing country-level data on children in institutions, the RMSE calculation is sensitive to the choice of dataset. For this reason, we present the estimates with the smallest RMSE as one type of estimates, rather than as the best estimates.

We did not report the imputed values from the multiple imputation in the text because the pattern of missingness in the data required us to drop approximately 50% of data on children living in institutions when doing the multiple imputation (appendix pp 18–19). We reported multiple imputation and other estimates in the appendix (pp 29–52).

We used STATA 14 in all analyses.

Role of the funding source

The funder had no role in study design, data analysis, data interpretation, or writing of the report. The funder shared their mailing list of potential data sources with the study team who used it to appeal for data from those sources.

	Total number (millions)	Average prevalence
Largest estimates		
Full data	9·42	0·51%
Official data	9·12	0·48%
Median estimates		
Full data	5·37	0·35%
Official data	5·85	0·37%
Lowest estimates		
Full data	3·18	0·30%
Official data	4·10	0·33%
Estimates from negative binomial regressions (95% CI)		
Full data	7·52 (7·48–7·56)	0·39% (0·39–0·39)
Official data	7·42 (7·38–7·46)	0·37% (0·36–0·39)
Estimates from imputation strategies with the smallest RMSE		
Full data	4·21	0·31%
Official data	5·96	0·37%

Estimates of level and prevalence of children living in institutions by income and region in 2015. RMSE=root-mean-squared error.

Table: **Global number and average prevalence of children living in institutions in 2015**

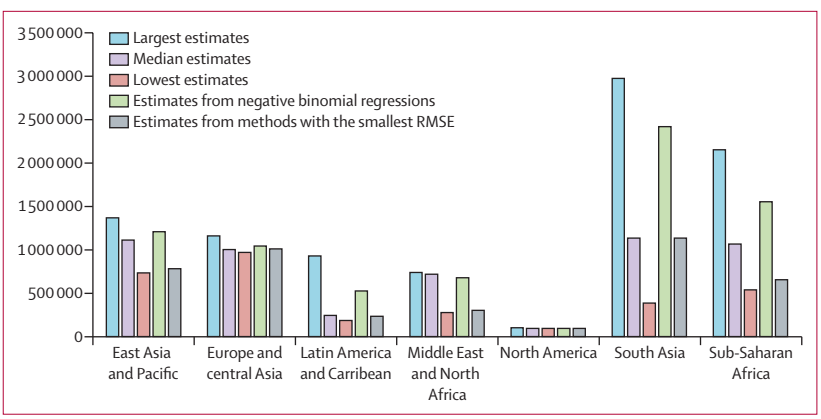

Figure 2: **Estimates of number of children living in institutions in 2015 derived from various imputation strategies, by region**
Estimates obtained with the use of the full data. RMSE=root-mean-square errors.

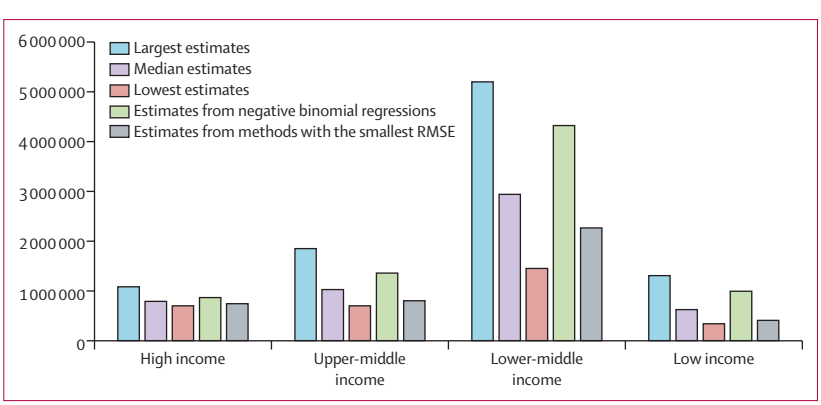

Figure 3: **Estimates of number of children living in institutions in 2015 derived from various imputation strategies, by World Bank income group**
Estimates obtained with the use of the full data. RMSE=root-mean-square errors.

The funder and the study team discussed the paper submission and potential journals for submission. All authors approved the final submission, and CD and CL made the final decision to submit the paper.

Results

Using national-level estimates, either existing or imputed, we generated estimates of average prevalence and total number of children living in institutional care in 191 countries in 2015 (table). We report here five estimates for the full data with the higher values and the official data with the higher values; the analysis of the lower-value data is presented in the appendix (pp 29–52).

Using the full data, the global-level estimates ranged between 3·18 million and 9·42 million, with a median of 5·37 million (table). The regression with negative binomial method yielded global estimates of 7·52 (95% CI 7·48–7·56) million. The global estimate from the methods with the smallest RMSE was 4·21 million. Estimates derived from other methods can be found in the appendix (pp 29–34). When applying the UNICEF method to the official data, the estimated total number of children living in institutions in 2015 was 2·93 million (appendix p 34). The average prevalence in the 191 countries in 2015 ranged from 0·30% (with the lowest global estimate) to 0·51% (with the highest global estimate; table).

Consistent with global estimates of children living in institutions, estimates in regions such as south Asia, sub-Saharan Africa, and Latin America were highly sensitive to the choice of estimation methods and datasets. For example, when using the full data, estimates for south Asia ranged from 0·38 million to 2·97 million (figure 2). When using the method with the smallest RMSE, estimates for sub-Saharan Africa ranged from 0·65 million with full data (figure 2) to 1·38 million with official data (appendix p 47). Estimates for North America (0·09–0·10 million) and Europe and central Asia (0·93–1·16 million) were stable across estimation methods and datasets; North America consistently had the lowest estimates among the six regions. South Asia or sub-Saharan Africa had the highest estimates in most cases. For global comparisons, estimates from the full data with the smallest RMSE method showed that south Asia had the largest estimated number of children living in institutions (1·13 million), followed by Europe and central Asia (1·01 million), east Asia and Pacific (0·78 million), sub-Saharan Africa (0·65 million), Middle East and North Africa (0·30 million), Latin America and the Caribbean (0·23 million), and North America (0·09 million; figure 2).

Among the four World Bank income groups, the lower-middle-income group consistently had the highest estimates across estimation methods and datasets, although with large variations in the values. When using the full data, the estimate for the lower-middle-income group was 4·32 million with negative binomial regression, but 2·26 million using the methods with the smallest RMSE (figure 3). For low-income countries, the estimates also varied, but were smaller than those for the lower-middle-income group (figure 3; appendix p 48).

Across the six regions, the estimated prevalence of children living in institutions for sub-Saharan Africa were typically the lowest, whereas estimates for Europe and central Asia were the highest (appendix pp 49–50). Compared with other income groups, high-income countries had the highest estimated average prevalence, whereas low-income countries had the lowest (appendix pp 51–52). This result was consistent across datasets and estimation methods.

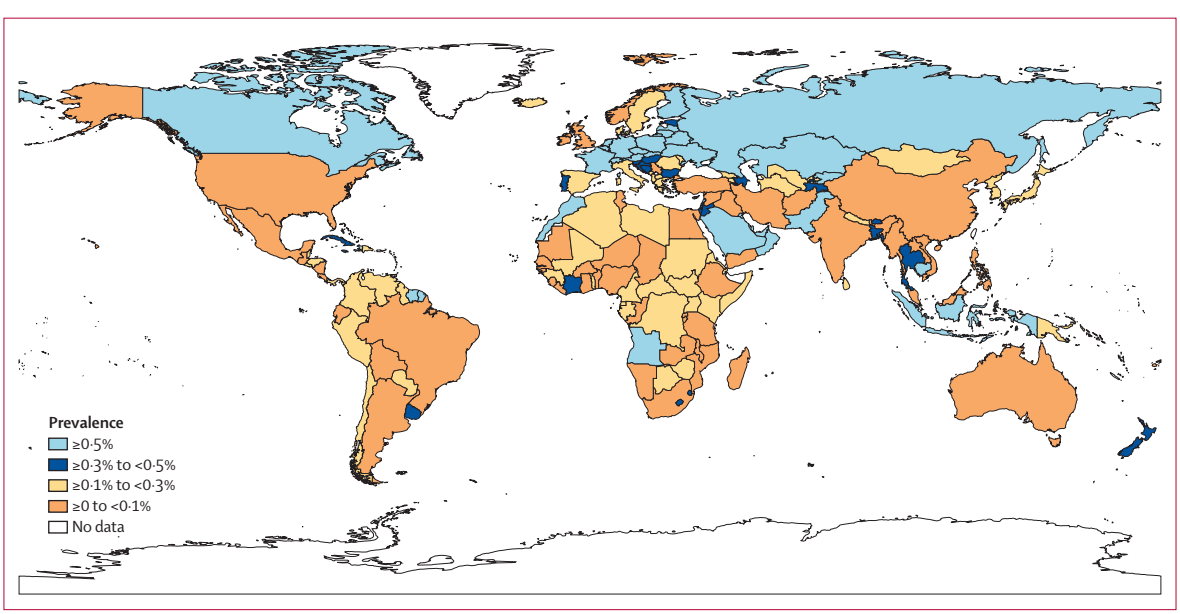

Figure 4: **Estimated prevalence of children living in institutions in 2015, by country**
Country-level estimates derived from the full data and by use of the imputation strategy with the smallest root-mean-square errors.

For countries without data in 2015, we observed a large variation in estimates across various estimation methods. In Nigeria, for example, estimates were very different from each other, from 3814 children (with the smallest RMSE method; appendix p 39) to 309 254 (with negative binomial regression).

For simplicity of presentation, we present here only the country-level estimates derived from the full data and the imputation strategy with the smallest RMSE. Country-level estimates derived from other methods are available on request. Of the 191 countries in 2015, 42 countries had an estimated prevalence of children living in institutions of 0·5% or higher and only one country (Haiti) was in the low-income group (figure 4). The five countries with the highest estimated prevalence were all from the high-income or middle-income groups and mostly in Europe and central Asia: Suriname (1·76%), Lebanon (1·45%), Ukraine (1·27%), Czech Republic (1·24%), and Slovakia (1·12%; appendix pp 41–46). 29 countries had estimated prevalence of 0·3% or higher and lower than 0·5%, 65 countries had estimated prevalence of 0·1% or higher and lower than 0·3%, and 55 countries had estimated prevalence lower than 0·1%. Country-level prevalences using the full data with lower values are shown in the appendix (pp 41–46).

The ten countries with the highest estimates of the number of children living in institutions in 2015 were Pakistan (0·543 million), Indonesia (0·458 million), India (0·337 million), Russia (0·263 million), Bangladesh (0·218 million), France (0·150 million), Democratic Republic of the Congo (0·120 million), Angola (0·109 million), Ukraine (0·100 million), and Morocco (0·097 million; appendix pp 35–40). The sum of children living in institutions in these ten countries accounted for approximately 57% of global estimates in 2015 (figure 5). The ten countries with the lowest estimates of children living in institutions are those with small child populations, with all estimates lower than 80 children (Andorra, Dominica, The Gambia, Libya, Mauritania, Monaco, Nauru, Palau, San Marino, and Tuvalu). Country-level numbers of children living in institutions derived from all datasets are presented in the appendix (pp 35–40).

Ranking	Country	Children in institutions (%)	Number of children in institutions
1	Pakistan	0·698	
2	Indonesia	0·537	
3	India	0·075	
4	Russia	0·951	
5	Bangladesh	0·38	
6	France	1·055	
7	DR Congo	0·296	
8	Angola	0·800	
9	Ukraine	1·265	
10	Morocco	0·869	

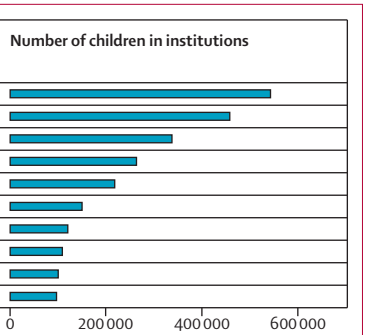

Figure 5: **Top ten countries with the largest number of children living in institutions in 2015**
Data obtained from the full data and by use of imputation strategies with the smallest root-mean-square errors.

Discussion

To our knowledge, we have constructed the most comprehensive dataset on the number of children living in institutions. By doing so and using various imputation strategies for countries with missing data, we reported here a range of estimates of the number of children living in institutions in 2015. The global estimates were highly sensitive to the methods and data used, ranging from 3·18 million to 9·42 million children, with a median of 5·37 million. When selecting the methods with the smallest RMSE or negative binomial regression with bootstrapping, this range narrowed to 4·21 million to 7·52 million. Regional estimates in south Asia, sub-Saharan Africa, and Latin America varied more than estimates for other regions when switching between estimation methods. Estimates in North America and Europe and central Asia were stable across estimation methods and datasets. Among the six regions, North America consistently had the lowest estimates of number of children living in institutions, whereas south Asia and sub-Saharan Africa had the highest estimates in most cases. Among the four income groups, lower-middle-income groups consistently had the highest estimates of number of children living in institutions, high-income countries had the highest estimated average prevalence of children living in institutions, and low-income countries had the lowest average prevalence. These results remained consistent when using different datasets and estimation methods.

Our global estimates are substantially higher than the 2·7 million estimate for 2012, produced by UNICEF.[16] This is probably a result of differences in both methods and data. Our selection of the method for imputing missing data had a sizeable influence on the estimates. If we adopted the UNICEF method and applied it to our official data, we would obtain an estimate of 2·93 million children, close to the UNICEF estimate. When applying the strategy with the smallest RMSE to the official data, the estimate was substantially higher, at 5·96 million. If we used the regression prediction with negative binomial model and bootstrapping method, the estimate would increase to 7·42 (95% CI 7·38–7·46) million.

Official data have been asserted to often bias downwards.[27] In countries where we had both an official estimate and the results of a primary data collection exercise, a comparison of the two most often supported this view (appendix pp 52–57). However, when using estimates from methods with the smallest RMSE, the estimates produced from the full data were approximately 1·75 million short of the estimates derived from the official data. This discrepancy is because more countries in the official data had no data or only one data point available and, as a result, prevalence was estimated, leading to larger values. For example, when using full

data, Mexico had two data points. According the estimation method with the smallest RMSE, its 2015 estimate should be obtained from linear extrapolation (27 717 children; appendix p 38). When using official data, the two data points could not be used because they were not officially reported. We needed to use a generalised linear model with binomial distribution to impute the missing value to obtain estimates with the smallest RMSE. This estimate for Mexico was 188 246 (95% CI 185 853–190 639), much larger than the existing non-official data.

The use of existing data for estimation has several limitations. First, our inclusion criteria for estimates of institutional care allowed for variability in definitions between countries. As a result, the definition used in this study is not well specified. Forms of care defined as institutions in one setting might have been excluded in another. We urge caution when citing the estimates or making comparisons with future estimations if data collected in the future have different definitions. Second, while we used all available data, the majority (65%) of data are still from official sources. As discussed in previous studies, official data might be underreported, especially in contexts where informal institutions are common.[16,28] Third, our estimates are sensitive to the choice of estimation methods and datasets. Fourth, our search strategy is likely to have missed some of the available data. Searches were done in English (complemented by a targeted search in Chinese for data on China and Korean for data on South Korea). Although this strategy yielded summary documents that drew on literature or reports in other languages, some data will only have been available in the original language. Similarly, we did not search non-English language government websites for official data. Finally, given the data available, we were only able to look at total numbers of children. We could not examine the breakdown by age, sex, and disability status. Similarly, we cannot comment on how many children are in each type of institution and on the variations in quality of care between institution types.

These limitations in estimation reinforce our primary conclusion: that considerable, and unacceptable, uncertainty exists regarding the number of children living in institutions. Agreement on a standard definition of institutional care is needed, possibly with subcategories allowing for the application of local definitions. This would allow a baseline to be set and progress to be monitored. Reporting of disaggregated data by age, sex, disability, and type of institution, should be encouraged, particularly by official sources. Ultimately, efforts to monitor, report on, and respond to children in institutional care should be government led. Recent work in Ghana, by its Ministry of Gender, Children, and Social Protection with support from the US Agency for International Development and UNICEF, provides a useful example. They have undertaken mapping exercises, established new standards, and are seeking to encourage alternatives, such as foster care.[29]

In the interim, we need to assess the validity of official statistics, especially for countries with large child populations that report very low numbers of children in institutions. We need to investigate differences between regions and country income groups to improve our understanding of the pattern of institutionalisation.

Even with the existing uncertainty, what is nonetheless clear is that millions of children live in institutions, despite the known risks to their health and development. Reducing the rate of institutionalisation through preventive interventions is possible.[2] Deinstitutionalisation can be risky if children are rushed to inappropriate alternatives or back into unsafe home environments but, with careful planning, most children in institutions can be reunited with their families or placed in appropriate foster care arrangements, and the minority for whom this is not possible can be placed in small high-quality group homes able to provide specialised care.[1,25] This might require an upfront investment in prevention and deinstitutionalisation, but is likely to be cost saving in the long run.[1] The high costs of institutions would be avoided, and the long-term health, productivity, and wellbeing of children would be improved.[1]

Contributors

CD and CL designed the study and wrote the paper. AS, KW, and JH did literature reviews and compiled the dataset. CL led the data analysis. CD and CL did data interpretation. All authors reviewed the drafts, made critical comments, and approved the final submission.

Declaration of interests

CD and CL report grants from the Lumos Foundation during the conduct of the study. All other authors declare no competing interests.

Acknowledgments

Funding for this study, the preparation of this paper, and the corresponding report was provided by the Lumos Foundation. We acknowledge the assistance of the Harvard Countway Library in the design of the systematic review of published literature, the Better Care Network and Eurochild for sharing contacts to support the review of the grey literature, and to The Global Study for Children Deprived of Liberty, for sharing their data. Chrissie Gale and Florence Martin provided valuable advice in the design phase of the study. We thank David L Liu for his invaluable contribution to the literature review, Dr Gunther Fink for a helpful review, and Prof Linda Richter for her astute guidance. We are grateful to the Lumos Foundation for their support, in particular Chris Cuthbert and Alex Christopoulos.

Editorial note: the *Lancet* Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

References

1 Berens AE, Nelson CA. The science of early adversity: is there a role for large institutions in the care of vulnerable children? *Lancet* 2015; **386:** 388–98.

2 Dozier M, Zeanah CH, Wallin AR, Shauffer C. Institutional care for young children: review of literature and policy implications. *Soc Issues Policy Rev* 2012; **6:** 1–25.

3 Csaky C. Keeping children out of harmful institutions: why we should be investing in family-based care. 2009. Save the Children UK. https://resourcecentre.savethechildren.net/node/1398/pdf/1398.pdf (accessed Oct 10, 2018).

4 Browne E. Children in care institutions. K4D Helpdesk Report. 2017. Institute of Development Studies. https://assets.publishing.service.gov.uk/media/5b9a43caed915d666f681e10/029_Children_in_Care_Institutions_v2.pdf (accessed Oct 10, 2018).

5 UNICEF. Progress for children: a report card on child protection, No. 8. 2009. https://www.unicef.org/publications/files/Progress_for_Children-No.8_EN_081309.pdf (accessed May 1, 2018).

6 Jozefiak T, Sønnichsen Kayed N. Self- and proxy reports of quality of life among adolescents living in residential youth care compared to adolescents in the general population and mental health services. *Health Qual Life Outcomes* 2015; **13:** 1–12.

7 Nelson CA, Fox NA, Zeanah CH. Romania's abandoned children: deprivation, brain development and the struggle for recovery. Cambridge: Harvard University Press, 2014.

8 Sheridan MA, Fox NA, Zeanah CH, McLaughlin KA, Nelson CA. Variation in neural development as a result of exposure to institutionalization early in childhood. *Proc Natl Acad Sci USA* 2012; **109:** 12927–32.

9 Hermenau K, Hecker T, Elbert T, Ruf-Leuschner M. Maltreatment and mental health in institutional care—comparing early and late institutionalized children in Tanzania. *Infant Mental Health J* 2014; **35:** 102–10.

10 McCall RB. The consequences of early institutionalization: can institutions be improved?—should they? *Child Adolesc Ment Health* 2013; **18:** 12025.

11 van Ijzendoorn MH, Luijk MPCM, Juffer F. IQ of children growing up in children's homes: a meta-analysis on IQ delays in orphanages. *Merrill-Palmer Q* 2008; **54:** 341–66.

12 Pollak SD, Nelson CA, Schlaak MF, et al. Neurodevelopmental effects of early deprivation in postinstitutionalized children. *Child Dev* 2010; **81:** 224–36.

13 Whetten K, Ostermann J, Whetten RA, et al. A comparison of the wellbeing of orphans and abandoned children ages 6–12 in institutional and community-based care settings in 5 less wealthy nations. *PLoS One* 2009; **4:** e8169.

14 Johnson DE, Gunnar MR. Growth failure in institutionalized children. *Monogr Soc Res Child Dev* 2011; **76:** 92–126.

15 Williamson J, Greenberg A. Families, not orphanages. Better Care Network working paper. 2010. https://bettercarenetwork.org/sites/default/files/Families%20Not%20Orphanages_0.pdf (accessed Oct 19, 2018).

16 Petrowoski N, Cappa C, Gross P. Estimating the number of children in formal alternative care: challenges and results. *Child Abuse Negl* 2017; **70:** 388–98.

17 NGO Working Group on Children without Parental Care in Geneva. Identifying basic characteristics of formal alternative care settings for children: a discussion paper. 2013. https://resourcecentre.savethechildren.net/node/7589/pdf/formal_care_settings_characteristics_march_2013_final1.pdf (accessed Jan 31, 2018).

18 UN General Assembly. Guidelines for the alternative care of children. UN Document A/RES/64/142. 2010. https://www.unicef.org/protection/alternative_care_Guidelines-English.pdf (accessed Dec 1, 2018).

19 International Household Survey Network. Central survey catalog. http://www.ihsn.org/survey-catalogs (accessed Oct 15, 2018).

20 The World Bank. World Bank country and lending groups. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (accessed Oct 15, 2018).

21 Baum CF. Modeling proportions. *Stata J* 2008; **2:** 299–303.

22 University of California Los Angeles Statistical Consulting Group. Regression models with count data. https://stats.idre.ucla.edu/stata/seminars/regression-models-with-count-data/ (accessed Nov 10, 2019).

23 Greene WH. Econometric Analysis, 5th edn. Upper Saddle River, NJ: Prentice-Hall, 2003: 740–52.

24 Kulesa A, Krzywinski M, Blainey P, Altman N. Sampling distributions and the bootstrap. *Nat Methods* 2015; **12:** 477–78.

25 Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast* 2006; **22:** 679–88.

26 McCoy DC, Peet ED, Ezzati M, et al. Early childhood developmental status in low-and middle-income countries: national, regional, and global prevalence estimates using predictive modeling. *PLoS Med* 2016; **13:** e1002034.

27 UNICEF. Progress for children: a report card on child protection. 2009. https://www.unicef.org/publications/files/Progress_for_Children-No.8_EN_081309.pdf (accessed Aug 25, 2019).w28 Mulheir G, Browne K. De-institutionalising and transforming children's services: a guide to good practice. 2007. https://resourcecentre.savethechildren.net/node/5995/pdf/5995.pdf (accessed Aug 25, 2019).

28 Mulheir G, Browne K. De-institutionalising and transforming children's services: a guide to good practice. 2007. https://resourcecentre.savethechildren.net/node/5995/pdf/5995.pdf (accessed Aug 25, 2019).

29 Ministry of Gender, Children and Social Protection. Mapping of residential homes for children in Ghana. 2018. https://www.unicef.org/ghana/media/1716/file/Mapping%20of%20residential%20homes%20in%20Ghana.pdf (accessed Oct 14, 2019).