**Child health research and planning in Europe disadvantaged by major gaps and disparities in published statistics**

Michael J Rigby 1 \*, Shalmali Deshpande 2, Mitch E Blair 2

1. Schools of Social, Political and Global Studies, and Primary, Community and Social Care, Keele University, UK

2. Section of Primary Care and Public Health, Imperial College London

*\*Corresponding Author*

Lavender Hill

6, Carrighill Lower

Calverstown

Kilcullen

Co. Kildare

R56 DT91

Ireland

m.j.rigby@keele.ac.uk

Tel: 00353 87765 3399

**Child health research and planning in Europe disadvantaged by major gaps and disparities in published statistics**

Michael Rigby, Shalmali Deshpande, Mitch Blair

# Abstract

Background. Population data such as mortality and morbidity statistics are essential for many reasons, including giving context for research, supporting action on health determinants, formulation of evidence-based policy for health care, and outcome evaluation. However, when considering children, it is difficult to find such data, despite children comprising one fifth of the European population and being in a key formative life stage and dependent on societal support. Moreover, it would be expected that there should be confidence in the key child health data available, with little to no discrepancy between recognised health statistic databases.

Methods. This study explored the main health databases in or including Europe to collate child mortality data, for both all-cause and specific cause mortality. Tables were constructed for comparison of values and rankings.

**Results**. The results show that there are major differences in reported mortality data between two prominent health statistic databases, difference in coding systems, and unannounced changes within one of the databases.

**Conclusion.** The lack of health data for children seems compounded by challenges to the trust and credibility which are vital if these data are to have utility. Children and society are the losers, and resolution is needed as a priority.

198 of Max 250 words

# Key Words

Child health; published data; evidence; policy; evaluation

**Child health research and planning in Europe disadvantaged by major gaps and disparities in published statistics**

# Introduction

Research into and planning of children’s services require robust baseline data. Not only are children a dependent population group, for whom society has a particular duty of provision and care, but they are a dynamic group. The individual child matures rapidly, but additionally health conditions, child and parental attitudes to services, and determinants of health, also change fast. This puts a particular focus on good research, evaluation, and planning.

Yet set against this need, important baseline data are somewhat sparse, children being not well represented in most types of health data, from morbidity to health behaviour, in part due to challenges of data capture. This results in children are not being well represented in health data sources in Europe [1].

The Models of Child Health Appraised (MOCHA) project had a particular interest in this topic given its remit of assessing models of primary health care in the then 30 European Union (EU) and European Economic Area (EEA) countries [2], funded under the European Commission’s Horizon 2020 programme from 2015 to 2018. Core to the MOCHA project was the focus on the child – defined according to the United National Convention of the Rights of the Child (UNCRC) as a person from birth up to their 18th birthday [3]. All EU and EEA countries are signatories to this convention.

In order to undertake the function of appraisal, one key aspect was to relate population and epidemiological data for the study countries to various aspects of health service provision, from spend levels to workforce numbers and training, in order to consider optimal outcomes for the type and volume of input and the population needs to be addressed. Ascertaining comparable baseline data for all 30 countries was clearly a core requirement, and recourse was made to the reputable published public health datasets. The intention was to seek to relate comparative position in the distribution of core child health status in each of the 30 countries with aspects of service structure and delivery, in order to identify correlations, and if possible indicate putative causality. In this need for baseline data the MOCHA project was no different from any other child health research or policy-setting activity, other than that it had a focus of expertise and resources. However, a major problem was identified, as reported here, and this will compromise other child health studies and service design and planning.

# Methods

A central initial objective of the MOCHA project was to create tables of child population, cause-related child mortality rates using broad groupings of diagnoses, and if possible some key morbidity or service use data. These were planned to drive deeper focussed studies, and power emerging major conclusions; the same needs would apply for many other researchers or policy-makers

## Children are not counted

The first problem is the variation in how children are counted. While the neatness of five-year age groupings is understandable, it cuts across any attempt at serious consideration of children’s health needs or outcomes by the legal definition enshrined also in the United Nations Convention on the Rights of the Child (UNCRC) [3], namely persons from birth up to their 18th. birthday. Perforce either 0-15 years has to be adopted as the total child population when consulting databases, discounting the interests of adolescents at a key stage in their lives, or alternatively 0-20 years is adopted, meaning that 18 and 19-year-old adults have to be included though their service provision and many needs are different, as of course is their legal status.

Moreover, the United Nations Committee on the Rights of the Child oversees a process of countries individually reporting on their implementation of the Convention. Sadly, Eurostat [4] and the WHO European Health for All Database [5] fail totally, as they do not have a category matching children 0-17 years inclusive, and therefore cannot monitor all children or factors affecting them, nor aid others in their own monitoring functions. The Institute for Health Metrics and Evaluation (IHME), creators of the Global Burden of Disease (GBD) calculations [6], would fail too but they have the ignominious position of being located in the one significant country globally that has not ratified the UNCRC, namely the United States of America.

## Mortality

Even if counting children is not achieved by the major statistical teams, hopefully calculating mortality rates – even for rigid five-year age bands – should not be difficult. Developed countries have strict and presumed reliable means of registering, reporting and aggregating deaths. Every EU country also has quinquennial or decennial census counts which establishes the number of children on census years, and processes for making year-on-year adjustments for change due to inward and outward movement as well as death so as to create population values for inter-censal years. Hence both the numerator (deaths) and the denominator (resident children) should be available to enable calculation of annual mortality rates which can be trusted as reliable. However, this proves not to be the case.

In order to progress its work, researchers in the MOCHA project sought to produce comparable data on overall, and selected specific, causes of death for each country in comparable format. Though it has many valuable statistics, the WHO Health for All database does not compile this data. However, Eurostat and IHME-GBD do produce it and project researchers consulted both the Eurostat and the IHME-GBD data sources to ensure robustness. What was not expected was the difference in published data, including not only death rates but also the ranking of countries, between the two sites, and secondly that one of the published data tabulations would prove not to be stable.

# Results

First, the basic question of child mortality. The authors have already reported in a brief communication elsewhere, the discrepancy between Eurostat and IHME data for death rates for children aged 10-14 years, this age-group being taken as a sample less affected by post-partum sequalae and infectious diseases, and also adolescent risk-based accidents [7]. That analysis showed important differences in individual countries between Eurostat and GBD data as of August 2018. Secondly, the authors subsequently identified a change in the data posted by the GBD study two months later, for the same mortality rate definition, and they published an analysis of the degree of change [8]. Death rates per country changed by between an upwards change of 29.28 % (Slovenia) to a decrease of 29.12 % (Ireland).

Table 1 now summates these two differences and changes into one table, and shows that after the GBD adjustments the difference between the two calculations of child mortality in this age group is between 29.8% upwards of the Eurostat rate (Hungary) and 38.5% lower (Iceland).

Table 1 about here

Figure 1 shows these three rates graphically. There is no obvious pattern as to which countries have strong concordance and which have considerable variation. The strongest concordance is with data for Portugal and Romania.

[Figure 1 about here]

It will be seen that there are noteworthy differences between the computed values, but also from the earlier publications [7, 8] that there is little match as to the rankings of the 30 countries between the two sources. Indeed, no countries have the same value and only two countries have the same rank order across the databases.

## Cause-Specific Mortality

Important for assessment of needs and evaluation of services is cause-related mortality. A first topic studied within the MOCHA project was mortality in older children excluding accidents. This is important as accidents are broadly outside the direct scope of health services, being primarily related to road traffic events, accidental drowning, activity-related events, and inter-personal violence. These are indeed amenable to public health and societal programmes, as shown for instance by the Child Safety Action Plan project [9], rather than by primary health care services. By excluding accidental deaths, analysis was intended to show deaths caused by illness and other health-related conditions.

As adolescent deaths have overtaken under-fives mortality in most countries in terms of policy emphasis, and survival of those with complex medical conditions has improved, for this aspect of the data availability review the researchers chose to focus on 15-19 year old age group, fully recognising that this spans beyond the usually defined period of childhood. They hypothesised that if accidental and violent deaths were excluded, deaths in this age-group could be considered an outcome of exposure to the health care system in the preceding childhood years as the main influence. Figure 2 shows the variation between Eurostat and IHME-GBD data, of children whose cause of death is not related to accident or violence.

[Figure 2 about here]

This figure shows the very great differences in this mortality rate for each country, as compiled by the two databases. With variation of this magnitude, and unable as end users to have the ability to drill down into source data and contributing calculations, this data set was too challenging in its positioning of each country to be of any value for the core purpose of seeking correlation with patterns of resource or health service provision structure. Only rates can be shown, not actual counts, as Eurostat does not publish counts, while the GBD method of estimation shows numbers of deaths with decimal points, which is difficult for many non-statistician users to accept as credible.

Intrigued by this disparity, the researchers wondered if it applied to all causes of childhood death. Fine granularity comparison by cause is difficult as the two databases use different systems of coding death – WHO and Eurostat use the International Classification of Disease (ICD), while the Global Burden of Disease uses a customised four level hierarchy of causes [1]. The researchers therefore selected death by drowning (defined as accidental drowning and submersion) for analysis, as the coding allocations are similar, and at the stage of reporting of each death the diagnosis should not be in doubt in the majority of cases. The results are presented the results in Figure 3.

[Figure 3 about here]

For four small countries (identified by an asterisk by the name) Eurostat does not compile rates. The results show considerable variations in country-specific rates, though in broad terms the rankings are similar. However, the country differences are considerable. The GBD data are different from Eurostat by between 59% lower (Slovakia) and 475% higher (Spain). Indeed, in only five of the 20 countries with non nil data from both sources is the difference less than 25% of the Eurostat figure. Thus, with a very specific and clear-cut cause of death, there is considerable variation between Eurostat and the GBD. Thus for comparison at the condition-specific level, and from this of services targeted for specific conditions, cross-site validation of mortality data is severely challenged by both different coding systems, and different data where comparison appears possible.

# Discussion

Evidence is vital to underpin health planning, health policy, and investment in health, emerging out of the fundamental principles of evidence-based medicine [10], into recognition of its own value, given that health policy is important and thus it too should be evidence-based [11,12]. Morrato *et al* have emphasised the value of population-based data in providing a robust under-pinning for policy [13]. This was part of the approach taken by the MOCHA project, seeking to utilise published population data as one of the projects key analytic inputs.

It was already known that morbidity data, though in many ways potentially more valuable, would be hard to come by – primary care data in statistical form are hard to obtain though this project tried [14-16], while secondary care data are seldom available by small age group bands. Thus, the preferred approach – assessment of prevalence represented by children receiving treatment - is not to available. However, it would seem reasonable to be able to rely on recognised sources of mortality data as a broad guide, not least given the robustness of certification systems in developed countries and the harmonisation activities of Eurostat.

The issues presented here are not just a hindrance to the analyses of one large strategic study. More fundamentally worrying is that these variations will undermine not only credibility in the sources, but in evidence-based policy itself, over a much wider range of uses. If policy makers, including politicians, and in turn the press, are able to undermine well-intentioned plans on the basis that even the database calculations of relative size of problems do not agree, the adverse effect on grounded policy making and investment is serious. At face value these sources and their definitions are the same, and the point about such data publication is that they can be trusted as they read, and that detailed review of compilation algorithms is not necessary at the point of use. The end user should be able to trust what they read as being what the display claims to purport.

It is recognised that the Global Burden of Disease initiative is breaking new ground in important ways, and now has the active support of the World Health Organisation [17], but this needs to be based on a position of credibility. A recent GBD study on the global impact of alcohol consumption [18] has led to comment from a very senior statistician about the opacity of some of the calculations [19]. At the same time, and quite fairly, members of the IHME team have explained the need to estimate when looking at complex public health topics [20]. However, counting children’s deaths in European countries should not need significant estimation.

Following publication of the communication about changes in IHME-GBD death rates [8] the IHME-GBD team have published an explanation that their methods are constantly improving, and therefore they are able to refine earlier estimates, and feel a duty to do so in order to make better data available for use [21]. However, this good intention does nothing to create trust in publication of statistics in that or indeed any other source, and it does not address the core issue of research or policy formation based on those data subsequently being rendered arguably erroneous. This has the risk of playing into the hands of ‘post-truth’ lobbies who argue against scientific methods, and in favour of strong decisions based on ungrounded belief models.

Eurostat sources its material from European Union member state official statistics sources. These may indeed contain inbuilt error, while at the same time the resident population calculation needs some estimation – of errors in each country’s last census, and of population change in subsequent years. However, that would not seem to support the degree of difference including difference in ranking shown in these sample analyses.

Robust and trustworthy population data, including mortality data by main causes, can be considered vital to support ongoing health system evaluation and development. Understanding children’s statistics would seem vital, given all European nations’ ratification of the UN Convention on the Rights of the Child. Our findings suggest that significant work needs to be done to reconcile data differences. Now that WHO has formalised an agreement with the IHME-GBD and therefore by that action indicated trust in the data processes, it too has an interest in this conundrum. The world needs credible and reliable sources of core child health and mortality data.

2,506 of Max 3,000 words

# Key Points

* Recognised sources of European data do not match on key child population data, such as crude mortality rate.
* Both data values and rankings differ between Eurostat and IHME-GBD, with only a few exceptions; use of such data to appraise and compare national primary care systems is problematic.
* There is a core issue of whether to include estimation; this needs high level moderation and agreement so as not to critically underpin trust in data sources, research, and health planning.
* Credibility of key major data sources in the eyes of policy makers and public representatives will be diminished if there are major discrepancies of apparently hard facts such as total death rate for child age groups.
* Public health and health services research, already impeded by lack of good morbidity data, are further handicapped if not even mortality data can be used convincingly.

# Conflict of Interest

None.

# Funding

The Models of Child Health Appraised (MOCHA) project was funded by the European Commission through the Horizon 2020 Framework under the grant agreement number: 634201, but this work has been produced subsequently without funding. The sole responsibility for the content of this paper lies with the authors. It does not necessarily reflect the opinion of the European Union, or of the full project. The European Commission is not responsible for any use that may be made of the information contained therein.

# References

[1] Rigby M, *et al*. The Invisibility of Children in Data Systems; in Mitch Blair, Michael Rigby, Denise Alexander. Issues and Opportunities in Primary Health Care for Children in Europe: The Final Summarised Results of the Models of Child Health Appraised (MOCHA) Project; Emerald Publishing, Bingley, 2019; Open Access at https://www.emerald.com/insight/content/doi/10.1108/978-1-78973-351-820191011/full/html

[2] Models of Child Health Appraised: [www.childhealthservicemodels.eu](http://www.childhealthservicemodels.eu). Accessed [07/09/2018].

[3] Convention on the Rights of the Child: <https://treaties.un.org/doc/Treaties/1990/09/19900902%2003-14%20AM/Ch_IV_11p.pdf>. Accessed [07/09/2018].

[4] European Commission, Eurostat database: <https://ec.europa.eu/eurostat/data/database>. Accessed [10/09/2018].

[5] World Health Organisation, Health For All Database: <https://gateway.euro.who.int/en/datasets/european-health-for-all-database/>. Accessed [10/09/2018].

[6] Institute for Health Metrics and Evaluation, Global Health Data Exchange, Global Burden of Disease Results Tool: <http://ghdx.healthdata.org/gbd-results-tool>. Accessed [10/09/2018].

[7] Rigby M, Deshpande S, Blair M. Credibility in published data sources; Lancet, Vol 393, January 19, 2019

[8] Rigby M, Deshpande S, Blair M. Another blow to credibility in published data sources; The Lancet, Vol 394, 26, July 6, 2019.

[9] <https://www.childsafetyeurope.org/actionplans/index.html> (accessed 16 September 2019)

[10] Sackett D, Rosenberg W, Gray J, Haynes R, Richardson S. Evidence based medicine: what it is and what it isn’t. BMJ 1996;312:71–72.

[11] Solesbury W. Evidence Based Policy: Whence It Came and Where It's Going. Queen Mary University of London: ESRC UK Centre for Evidence Based Policy and Practice 2001.

[12] Dobrow MJ, Goel V, Upshur REG. Evidence-based health policy: context and utilisation. Soc Sci Med 2004;58:207-17.

[13] Morrato EH, Elias M, Gericke CA. Using population-based routine data for evidence-based health policy decisions: lessons from three examples of setting and evaluating national health policy in Australia, the UK and the USA. J Public Health 2007;29:463-71.

[144] Liyanage H, Hoang U, Ferreira F, De Lusignan S. Report of Measures of Quality and Outcomes derived from large datasets: Models of Child Health Appraised; 2018. Report No.: D14 5.2.

[15] Liyanage H, Hoang U, Ferreira F, Alexander D, Rigby M, Blair M, De Lusignan S. Availability of Computerised Medical Record System Data to Compare Models of Child Health Care in Primary Care Across Europe. Stud Health Technol Inform 2017;244:8-12.

[16] Liyanage H, Shinneman S, Hoang U, Ferreira F, Alexander D, Rigby M, Blair M, de Lusignan S. Profiling Databases to Facilitate Comparison of Child Health Systems Across Europe Using Standardised Quality Markers. Stud Health Technol Inform 2018;247:61-65.

[17] Memorandum of Understanding between WHO and IHME, 2018: <http://www.healthdata.org/sites/default/files/files/WHO-IHME%20MoU_052218.pdf>. Accessed [11/09/2018].

[18] The GBD 2016 Alcohol Collaboration. Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2018;392:1015-35.

[19] Spiegelhalter D. The Risks of Alcohol (again). Medium 2018 Aug 24.

[20] Leach-Kemon K, Gall J. Why Estimate?: <http://www.healthdata.org/acting-data/why-estimate?utm_source=IHME+Updates&utm_campaign=c062742ac8-Weekly_Email_Sep_6_2018&utm_medium=email&utm_term=0_1790fa6746-c062742ac8-422568873>. Accessed [11/09/2018].

[21] Dicker D, Nguyen G, Lopez AD, Murray CJL. Another blow to credibility in published data sources - Author's reply.[;](https://www.ncbi.nlm.nih.gov/pubmed/31282360) Lancet. 2019 Jul 6;394(10192):27-28.