Developing Guidance Documentation for Assessing the Suitability of Mortality Data Capturing Drowning Deaths for Drowning Prevention Users

Dr. Aoife A. Kervick, Dan Ryan, and Dr. Kiran Sarma

Draft submitted: 29/03/2017

Table of Contents

1. Introduction 3

2. Mortality Data Characteristics 4

2.1 Isolation of drowning-specific data 4

2.2 Type of drowning data captured 4

3. Data Coverage and Completeness 5

3.1 Data coverage 5

3.2 Data completeness 5

4. Data Quality 6

4.1 Means of data collection/generation 6

4.2 Data entry protocol 6

5. Data Accessibility 7

5.1 Ease of access 8

5.2 Ease of use 8

6. Conclusion 8

References 10

1. Introduction

Low and Middle-Income Countries (LMICs) are subject to disproportionately high rates of drowning mortality (e.g. Hyder et al., 2008; Peden & McGee, 2010). In order to effectively address this global health issue, evidence-based interventions, tailored to the particular country, or region, are required (Craig, Dieppe, Macintyre, Michie, Nazareth, Petticrew, 2008; Mitis, Sethi & Racioppi, 2010; World Health Organisation, 2014). A key drowning prevention objective, recommended by the World Health Organisation (WHO), involves the development of national water safety plans (WHO, 2014). In order to develop such a plan, 'accurate, timely, inclusive' drowning data are needed, to first, assess, and second, generate awareness of, drowning rates and risk (WHO, 2014, p. 36). Such data may also provide a means of evaluating the impact of any implemented drowning prevention measures or interventions.

A series of multi-country mortality databases, which document global drowning-related data, are currently available online. These include the WHO Mortality Database[1], the Global Burden of Disease (GBD) study[2], the International Disaster Database (EM-DAT) [3], and INDEPTH Network[4]. While all of these databases provide data in relation to drowning deaths, they do so in a number of different ways. First, each database differs in what they classify and count/include as a drowning fatality[5]. Second, their level of coverage of global drowning fatalities, and the completeness of the data supplied by their range of countries covered, to the database, varies. Third, the quality of the data provided by each database differs, due to their unique sources, and validation processes employed. Last, the accessibility of drowning mortality data tends to be database-specific, in terms of whether this is provided to users for real-time download, involves an online request procedure, specialist software for access, to payment of fees for source data.

As such, depending on the specific needs and standards of the drowning prevention user, certain mortality databases will provide data that is more suitable for their work than others. Selecting suitable mortality data is of crucial importance to ensure that the most relevant and accurate data is extracted to inform drowning prevention campaigns/measures. Unsuitable data could lead to incorrect processing, and interpretation of output, and ultimately, ineffective drowning prevention measures. In order to mitigate against this, the current report seeks to provide accessible, user guidance on assessing the suitability of mortality data for drowning prevention users. It does so with broad reference to the 'Data Appraisal and Assessment Criteria' developed as Deliverable A1 of the current project.

---

[1] WHO Global Mortality Database:http://www.who.int/healthinfo/mortality_data/en/
[2] Global Burden of Disease Study: http://www.healthdata.org/gbd
[3] The International Disasters Database: http://www.emdat.be/
[4] INDEPTH Network: http://www.indepth-network.org/about-us
[5] Of note, EM-DAT does not specify cause of death. Rather, it documents all fatalities that occurred due to the specific disaster.

2. Mortality Data Characteristics

The kind of data documented within a mortality dataset is a key consideration for prospective users when assessing data suitability for drowning prevention measures. First, whether or not drowning-specific data are documented, and can be isolated for analysis, must be assessed. Following this, the user needs to reflect on their purpose for accessing the data (e.g. to develop a National Water Safety Plan, or a disaster-preparedness programme), and whether the mortality dataset under consideration records the type of drowning data needed to achieve this.

2.1 Isolation of drowning-specific data

The level of detail at which cause of death (COD) is recorded within mortality datasets, varies from database to database. Although the International Classification of Diseases (Version 10) for example, is the gold-standard means of classifying all COD data (De Coster et al., 2006; WHO, 2016), this is not employed for all mortality datasets. As such, prior to selecting mortality data for use, whether or not drowning-specific fatalities can be isolated and extracted for use, must be established. The EM-DAT for example, captures mortality data in relation to categorised global disasters, documenting deaths due to water-related disasters such as flash floods, tsunamis, to avalanches. The figures provided relate to *all* fatalities incurred as a result of the specific water-related disaster however, and not just those due to drowning. As such, EM-DAT mortality data are not suitable for a user seeking to develop drowning-specific prevention measures, as they encompass deaths irrespective to drowning, with no means of separating these. Other data, such as those supplied by the INDEPTH Network, or GBD, which allow selection of drowning data alone, may be more suitable.

2.2 Type of drowning data captured

The *type* of drowning a prospective user is interested in preventing will also influence how suitable certain mortality datasets are for their work, as this differs from database to database. Intentional drowning deaths (due to suicide, or assault), for example, or drowning fatalities due to water transport accidents, are solely captured by the WHO Mortality Database to date. Likewise, drowning fatalities that are unspecified/undetermined, are documented by the WHO Database alone. The majority of mortality datasets solely provide data for drowning deaths due to accidental/unintentional submersion (e.g. the INDEPTH Network, GBD). Depending on whether the user's definition of a drowning fatality encompasses a broader range of drowning deaths (such as intentional drowning fatalities), only select mortality data (i.e. that from the WHO) may be suitable. Similarly, if the user seeks to examine drowning deaths with regards to the mechanism of drowning (e.g. in water, or following a fall), and nature of the body of water (e.g. a swimming pool, or bathtub) the individual drowned in, the WHO Mortality database alone provides this information, and is most suitable for use.

Last, if the user is interested in preventing drowning fatalities due to global disasters, or extreme environmental conditions, the most suitable data available to do so is sourced from the EM-DAT, or WHO Mortality Database (which documents deaths due to, for example, floods, or cataclysmic storms). Neither of these datasets provide drowning-specific data within these categories however, which must be acknowledged in assessing overall suitability.

3. Data Coverage and Completeness

The extent to which the mortality data being assessed provides data for the country/region(s) the user is seeking to develop drowning prevention measures for (referred to as 'data coverage') is another key consideration when assessing mortality data suitability. In addition, the extent to which the data provided for their area of interest, captures all of the drowning deaths that occurred during a certain timeline (i.e. 'data completeness'), should inform suitability assessments. Both provide indications of the accuracy and representativeness of the drowning-related mortality data documented. Here, representativeness relates to the extent to which the data captured provides a correct depiction of the total population.

3.1 Data coverage

Different mortality datasets record drowning data from different parts of the world. In assessing data suitability for drowning prevention, a prospective user should seek to identify mortality data that, optimally, contains data which covers (that is, is sourced from) the area(s) of interest that they wish to develop drowning prevention measures for, and for whom corresponding population data is available. The majority of multi-country datasets (e.g. the WHO Mortality Database, the INDEPTH Network) include a detailed, online list of the countries/regions that provide them with mortality and population data, and can be searched by a prospective user to assess its suitability in this regard. Of note, preference should be given to mortality data derived from the area of interest where possible, rather than estimates based on surrounding areas, for optimally accurate and representative data. This will be identified on the corresponding website for the mortality data (e.g. the GBD website).

3.2 Data completeness

The extent to which all the deaths that occur in the country/region are documented by a civil registration system/data capture tool (i.e. 'completeness') should also inform suitability assessments. It should do so, in that, the more complete mortality data is, the more accurate, and better representative it is of the deaths occurring in the county/region of interest of the prospective user. As such, more complete mortality data should be given priority if considering multiple datasets for use. Certain mortality datasets (e.g. the WHO Mortality Database) include measures of this, calculated by statistical methods, however many do not (e.g. EM-DAT), or are estimate-based (GBD).

4. Data Quality

The quality of the mortality data provided by the source is a crucial consideration when assessing the suitability of mortality data for drowning prevention, or any stakeholder use. Of particular importance, the tools and methods used to collect/generate mortality data can be examined to inform suitability assessments, as these attest to the quality of the data documented. Similarly, the protocol surrounding the assessment of raw data incoming to (or being reviewed within) the database, against specified standards, is indicative of data quality, and suitability.

4.1 Means of data collection/generation

How mortality data are collected/generated is of key importance when considering its suitability for use. Here, a prospective user should, where possible, give preference to mortality data collected through standardised, established means (e.g. registered death certificates, or Verbal Autopsy tool and software, e.g. see WHO, 2012; 2016), which have optimally had COD confirmed by qualified medical personnel. The WHO Mortality Database, for example, is solely populated by medically-certified, nationally registered deaths, where COD has been verified by a trained, medical practitioner, in line with ICD-10 criteria. If non-official codes are used, WHO experts replace these with the most appropriate ICD-10 codes. Conversely, the GBD study for example, uses data estimates to provide global mortality coverage, which, although in line with best practice and employing advanced modeling techniques, is a quality limitation. The uncertainty intervals (or error margins) associated with any mortality data estimates should be acknowledged in any quality, and suitability, assessment for prospective use. Of note, EM-DAT data is obtained from a wide range of sources, including media reports, which are often problematic due to errors in reporting, and should be acknowledged when considering quality.

That key data variables for understanding drowning risk and rates (e.g. the age and gender of the deceased) are included and recorded in the mortality dataset is another important consideration when assessing the quality of data as suitable for prospective use. The EM-DAT for example, do not include demographic data in any datasets available for download and use from their website, and may only provide such information following a formal request. Other mortality datasets (e.g. from the WHO, GBD or INDEPTH Network) typically include such information, crucial for designing targeted drowning prevention measures. Whether a detailed data dictionary/glossary for the variables listed in the dataset is available, is another key consideration for a suitability assessment, to ensure that the variables provided are described in sufficient detail to be analysed and interpreted correctly.

4.2 Data entry protocol

Mortality data, provided by a source database with an established, transparent protocol leading up to data entry, should be sought out when considering the suitability of certain mortality data for use, as this attests to the overall quality of the

data produced/entered. The GBD, for example, provides a detailed protocol document[6], outlining, amongst other items, their procedure for data synthesis, the estimation process, and how any disagreements regarding estimated figures are resolved prior to data entry. Ideally, this protocol should incorporate the use of error-detecting software (or conducting statistical analyses to identify outlying errors) when entering/compiling data into a dataset/database to maximise accuracy, and mitigate human error. The INDEPTH Network for example, routinely assess microdata with software designed to detect errors, attesting to the higher quality of the data provided by the dataset. Likewise, it should also specify what percentage of deaths recorded as mortality data are unspecified/undetermined (referred to as 'dump' or garbage' COD codes), and ideally, whether they are redistributed to more accurate, meaningful codes for drowning prevention (improving the overall quality of the mortality data), as they are, for example, within the GBD study.

The period of time elapsed between a fatality occurring, and its entry/publication online as part of a mortality dataset (i.e. the data's 'timeliness') is another important consideration when assessing data suitability with regards to quality. Trends in fatalities (including drowning fatalities) change over time, and current figures are needed to best inform drowning prevention measures. The WHO for example, use nationally registered vital statistics for their Mortality Database, data which are typically only made available online approximately 2 years after a death has occurred. This can reflect country-specific delays associated with holding inquests, or the time required for processing on behalf of the country/region, and/or the WHO, for example. While this is important to ensure high quality data, depending on the requirements for current, timely data of the prospective user, this delay may need to be considered when assessing data suitability.

Similarly, the extent to which the data is (or is not) regularly reviewed, and/or updated, should inform suitability assessments, in the context of data quality. Accurate, up-to-date data is needed to ensure trends in drowning fatalities are correct, and can reliably inform prevention measures. The EM-DAT for example, is updated internally each day, with publically accessible information being updated every three months, once all data has been validated and cross-checked across their numerous sources. Likewise, the GBD carry out scheduled 'rounds', in which their entire time series back to 1990 are re-estimated, using all and newly-available data, with previous datasets being archived. Of note, this means that previously published GBD estimates are not comparable.

5. Data Accessibility

A final consideration when assessing the suitability of mortality data for drowning prevention users is how accessible it is, that is, how easy the data are to access, and

[6] http://www.healthdata.org/gbd/about/protocol

how easy they are to interact with, and use. Here, a prospective user should reflect on accessibility-related factors such as the timeframe in which they require the data, and the expertise of the individual(/s) intending to analyse it, to inform suitability assessments.

5.1 Ease of access

The majority of multi-country mortality datasets feature some form of data visualisation tool (e.g. the visualisation tools of the WHO Mortality Database), or dataset download options (e.g. the GBD Results Tool, which allows certain variables to be downloaded as a personal CSV datafile) to provide real-time access to basic, drowning-related data. Access to more detailed microdata however, typically required to inform a high-quality drowning intervention, can take much longer. To access microdata on water-related disasters from EM-DAT, for example, a prospective user must complete an online request procedure. Any such requests are reviewed on a case-by-case basis, and may take several weeks to process, and months before access is granted. Similarly, access to any published datasets provided by the INDEPTH Network requires the completion of a formal online application, evaluated during a similar, unspecified time period. Conversely, the WHO Mortality Database microdata are available for immediate download, and as such, may prove a more suitable data option, should a user need to work within a strict timeframe. In addition, access to certain types of data may also incur fees (the GBD website states that access fees may be charged by certain sources), which depending on the financial capabilities of the prospective user, may render the data unsuitable for use.

5.2 Ease of use

The need for the prospective user to possess technical expertise or even specialist software to work with drowning mortality data varies from dataset to dataset. The detailed, ICD-10 coded microdata available for download from the WHO Mortality Database for example, requires statistical processing software (such as SPSS, or Stata) to facilitate this, and an experienced researcher to work with the hundreds, to thousands, of data points provided. As such, depending on the competency of the user (and/or their research team) in working with high-level mortality data, more basic data from other sources (e.g. the GBD, GHDx tool) may prove more suitable, to ensure it can be used correctly, and with ease. This may also depend on the level of online guidance documentation provided regarding the data download process, which should be consulted to inform suitability assessments.

6. Conclusion

This report has outlined a series of key considerations for assessing mortality data suitability for users seeking to design drowning prevention measures. Depending on the specific data needs and standards of the particular user (e.g. they may accept estimate-based mortality data, whether others may not), the considerations outlined in

the present report should provide them with sufficient information to judge how suitable the data being assessed are, for their use.

In summary, the extent to which drowning-specific data can be isolated for analysis within a dataset needs to be examined, as do the different types of drowning captured (e.g. accidental, intentional, water transport, unspecified/undetermined, due to extreme environmental conditions etc.), depending on the kind of drowning prevention measures the user is seeking to design (e.g. disaster-preparedness). Whether the mortality data covers (i.e. includes data from) the country/region(s) that the user wishes to target with drowning prevention measures is another important factor when assessing mortality data suitability. Likewise, the extent to which the data for the country/region(s) of interest is complete (i.e. includes all deaths that occurred there) needs to be examined. Both of these will provide indicators of how accurate and representative the data is for use and accordingly how suitable they are for a prospective user.

Mortality data quality should also be an influential factor when assessing the suitability of mortality data for drowning prevention use. Data that have been collected using established instruments, ideally with a medically-certified COD provided, are of optimal quality and most suitable for use. Key variables for understanding drowning risk (e.g. aggregate data for age, gender etc.) should also be included, with a data dictionary/glossary provided for all variables. Data for which there is a detailed protocol for data entry (including whether error detecting software is used to enter data, and how 'dump' or 'garbage' codes of death are dealt with) should also be regarded as optimal here. The timeliness of the data (i.e. how quickly it is published), and how regularly it is reviewed/updated should also influence assessments of suitability for use.

Last, how easy the mortality data is to access and use is an important consideration for a prospective user. Depending on the timeframe in which they want to develop their drowning prevention measures, and their expertise (or the expertise of their team) in working with complex mortality data, this may influence the suitable mortality dataset they ultimately choose to access.

References

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal, 337,* A1655, 979-983. Doi: 10.1136/bmj.a1655.

De Coster, C., Quan, H., Finlayson, A., Gao, M., Halfon, P., Humphries, K. H., Johansen, H., Lix, L. M., Luthi, J.-C., Ma, J., Romano, P. S., Roos L., Sundararajan, V., Tu, J. V., Webster, G., & Ghali, W. A. (2006). Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: report from an international consortium. *BMC Health Services Research, 6(*1), 77. Doi: 10.1186/1472-6963-6-77.

Hyder, A. A., Borse, N. N., Blum, L., Khan, R., El Arifeen, S., & Baqui, A. H. (2008). Childhood drowning in low- and middle-income countries: Urgent need for intervention trials. *Journal of Pediatrics and Child Health, 44*(4), 221- 217. Doi: 10.1111/j.1440-1754.2007.01273.x.

Mitis, F., Sethi, D., & Racioppi, F. (2010). Implementation of evidence-based interventions to prevent drowning: results from a WHO collaborative project. *Injury Prevention, 16* (1), A224. http://dx.doi.org/10.1136/ip.2010.029215.797

Peden, M. M., & McGee, K. (2003). The epidemiology of drowning worldwide. *Injury Control and Safety Promotion, 10*(4), 195-199. Doi: 10.1076/icsp.10.4.195.16772.

World Health Organisation (2016). *Classifications.* Retrieved from: http://www.who.int/classifications/icd/en/

World Health Organisation (2016). *Verbal autopsy standards: The 2016 WHO verbal autopsy instrument.* Retrieved from: http://www.who.int/healthinfo/statistics/verbalautopsystandards/en/

World Health Organisation (2014). *Global report on drowning – Preventing a leading killer.* Retrieved from: http://www.who.int/violence_injury_prevention/global_report_drowning/en/

World Health Organisation (2012). *Verbal autopsy standards: The 2012 WHO verbal autopsy instrument.* Retrieved from: http://www.who.int/healthinfo/statistics/WHO_VA_2012_RC1_Instrument.pdf?ua=1