



Get
every one
in the picture

Data Sources

Workshop on data analysis and report writing for civil registration based vital statistics

Nadi, Fiji

30 January – 03 February 2023

**Bloomberg
Philanthropies**



Pacific
Community
Communauté
du Pacifique



UNITED NATIONS
ESCAP

Economic and Social Commission for Asia and the Pacific

Population and institution based sources

- Population-based data sources are those that are representative of the whole population:
 - Population surveys
 - Censuses
 - Civil registration
- Institution-based data sources are those collected routinely from administrative and operational activities:
 - Health Information Systems (HIS)
 - Hospital discharge data
 - Police records for attended deaths
 - Social security records
 - Health facility surveys, where data are collected in an institution

Routine vs. periodic collections

◆ Routinely collected administrative data

- ◆ CRVS
- ◆ HIS/Health Information System Data
- ◆ Health facility data
- ◆ Other: demographic surveillance sites

◆ Periodic collection

- ◆ Censuses
- ◆ Household Surveys

Routine civil registration

- ◆ **Civil registration** provides a legal basis for the recording of vital events such as live births & deaths
- ◆ In countries with well-functioning CRVS systems, it is a legal requirement that:
 - a medical practitioner completes a *death certificate* whenever somebody dies
 - vital events, such as births and deaths, are registered
- ◆ An efficient routine CRVS system, with medical certification of CoD, provides ongoing and relatively low-cost data collection and therefore timely mortality data for decision making

Considered the “**gold standard**” for birth and death data

HIS (Health Information System) Data

- ◆ Vital events such as births and deaths are usually recorded (and notified) through routine data collections within the health system.
- ◆ Health data collections are primarily to inform operational decisions, and CoD is central to this purpose.
- ◆ May include:
 - ◆ medical certificates
 - ◆ community nursing reports,
 - ◆ facility based data
- ◆ Can provide population denominator data
- ◆ **Important to consider if this is population-based data or facility-based data**

Census data

A population census is a compulsory, universal and simultaneous enumeration of the national population, conducted on a periodic basis.

- ◆ Primary source of mortality data in PICTs
- ◆ Can provide population denominator data and mortality data
- ◆ Information on births and deaths may be derived
 - ◆ directly
 - ◆ Indirectly
 - ◆using changes in population (by age and sex) based on survival and migration
 - ◆ reliant on accurate migration data

Using census data for fertility/mortality analysis



Direct measures

- Asks about births and deaths in the household over a nominated reference period
 - usually within the last 12 months.
- Sometimes takes a complete birth history (not usually used in the census)

Indirect measures

- Collects some information on births and deaths to approximate when these events would have occurred.
 - Children ever born – children surviving
 - Orphanhood
 - Widowhood
- These measures can then be used as inputs for model life tables that generate estimates of age-specific measures of mortality and LE.
- Models that use a single input parameter based on childhood mortality have been shown to have a tendency to under-estimate mortality in PICTs

Note: Although censuses provide nationally representative data which can be used in the estimation of fertility and mortality analyses, censuses are usually only conducted every 5-10 years, and so consideration should be made for their timeliness.

Periodic Household Surveys

Household surveys collect comprehensive and diverse socio-demographic data on a sample of households in a population.

- ◆ Examples include: Demographic and Health Surveys (DHS) and the UNICEF Multi-Indicator Cluster surveys (MICS)
- ◆ Not ongoing data collection, but are collected in a specific period of time
- ◆ Collect data for a proportion of the population considered to be representative of the broader population of interest
- ◆ Usually collect information on self-reported birth registration
- ◆ Can be used to produce direct/indirect estimates of infant and child mortality, based on birth history data
- ◆ The reliability of estimates from survey data is driven by how well the sample selection reflects the broader population of interest
- ◆ Surveys are also subject to
 - ◆ recall bias, and
 - ◆ response bias
 - ◆ Sampling errors

Demographic surveillance sites

- ◆ A demographic surveillance system captures all vital events in a specified area.
- ◆ Often combined with disease detection (sentinel surveillance)
- ◆ Used in PNG
- ◆ Unlikely to be a suitable solution in other PICTs with small populations



Registries

- ◆ Source of outcome data
 - ◆ i.e. – cancer registries, pregnancy registers
- ◆ May be population or institution based
 - ◆ Population registries much more common where there is national testing facilities
- ◆ For deaths: registries record anyone who dies with the disease rather than from the disease (as the designated underlying cause)
 - ◆ Therefore the data is not directly comparable to that obtained from a civil registration system
- ◆ Pregnancy registers often maintained at health facility level, but rarely electronic and generally not centrally collated
 - ◆ therefore the data is not directly comparable to that obtained from vital registration.

Comparison of data sources for mortality

Data Source		Periodicity	Sample frame	Period of interest	Data collection (mortality level data)	CoD Data Collected?
Census		Periodic – 5-10 years	Whole Population	Retrospective	Direct - (deaths in the household) Indirect – partial birth history (CEB/CS) & orphanhood data	No
Survey	DHS	Periodic – ~ 5 years	Selected sample – representative of whole population	Retrospective	Direct – complete birth history	No
	MICS	Periodic – ~ 5-10 years	2 stage clustered sample – representative of whole population	Retrospective	Indirect - partial birth history (CEB/CS)	No
	Other household based surveys	Usually once-off	Varies	Retrospective	Varies	Possible - using verbal autopsy
Routine vital registration	Civil Registration	Continuous	Whole population (depending on coverage)	Current	Direct reporting of event	Yes
	Health vital registration	Continuous	Whole population (depending on coverage)	Current	Direct reporting of event	Yes
	Hospital discharge records	Continuous	Hospital cases only	Current	Direct reporting of event	Yes
Other routine databases	Various	Continuous	Varies – usually targets sub-population of specific interest.	Current	Direct reporting of event	Usually limited
Demographic surveillance Sites		Continuous	Selected areas – usually not representative of whole population over time.	Current	Direct reporting of event	Yes

(source: Carter, 2013)

Which data sources will we be using?

- ◆ For the production of vital statistics during this workshop, we will be using birth and death data collected by either the [Civil Registration](#) office and/or the [Ministry of Health](#).
- ◆ We will also use population data derived from your most recent census or projections (developed nationally or internationally e.g., by SPC or UNWPP) as denominator data to compute various indicators.

Where does your data come from?

- ◆ Think about the data you brought to the course and where it comes from...
 - ◆ Why was it collected?
 - ◆ How was it collected and by whom?
 - ◆ What are the strengths and weaknesses of that data collection?
 - ◆ What impact is that likely to have on your data?