Pacific Community
Communauté du Pacifique

A review of key concepts: Coverage & completeness

Karen Carter – CRVS Specialist, Pacific Community

UNITED NATIONS EXPERT GROUP MEETING ON THE METHODOLOGY AND LESSONS LEARNED TO EVALUATE THE COMPLETENESS AND QUALITY OF VITAL STATISTICS DATA FROM CIVIL REGISTRATION
New York, 3-4 November 2016

Importance of evaluating CRVS data quality

- As recognition of the importance of Civil Registration and Vital Statistics (CRVS) systems in underpinning development progress has increased over the last few years, there has been increasing attention to understanding the quality of the data that is collected through these systems.
- For CRVS data on population and health outcomes to be useful to governments for both planning and monitoring interventions, we need both to understand the population which the data represents, and to be assured that the data is “fit for purpose”.
- At the national level, this means that the data is relatively complete and that it reliably represents the population in which we are interested.
- For identification purposes however, the “margin of error” which we should be prepared to tolerate in the collection of CRVS data, must by necessity be even smaller.

Key concepts - Coverage

- Coverage is a measure of the population that the registry system serves.

- This is primarily, although not exclusively, an issue of access to the reporting system- and may be influenced by geography or other considerations such as the legal intent of the system, social or cultural influences.

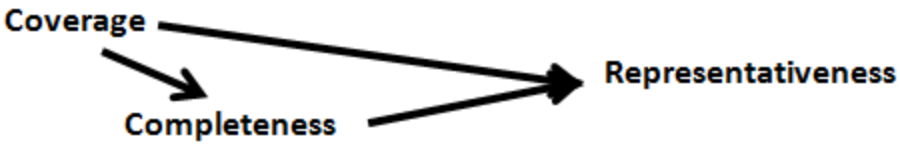
- 241: A basic requirement within a vital statistics system is that each vital event occurring within the geographical area covered by the system be registered once and only once for legal purposes and reported for statistical purposes within the time period stipulated by law, thus enabling 100 per cent—or universal—coverage
  - *UN Principles and Recommendations, 2014*

- The P&R also recommend that the “population of interest” should be the “usual resident population” where possible.

Key concepts - Completeness

- If coverage describes the population for which registration is actually possible, completeness is a measure of how well we actually capture all of the events in that population.
- Measuring completeness at a national level (as indicated in the SDG targets, and plans such as the Pacific Regional Action Framework) implicitly assumes national coverage as we cannot register events where there is no access to registration.
- It is this national (or population) measure of completeness that is however most of interest to statisticians as this is the figure that is the key factor in whether the registry data can be used to calculate representative vital statistics indicators at the population level.

Figure 1: Relationship between Coverage and Completeness

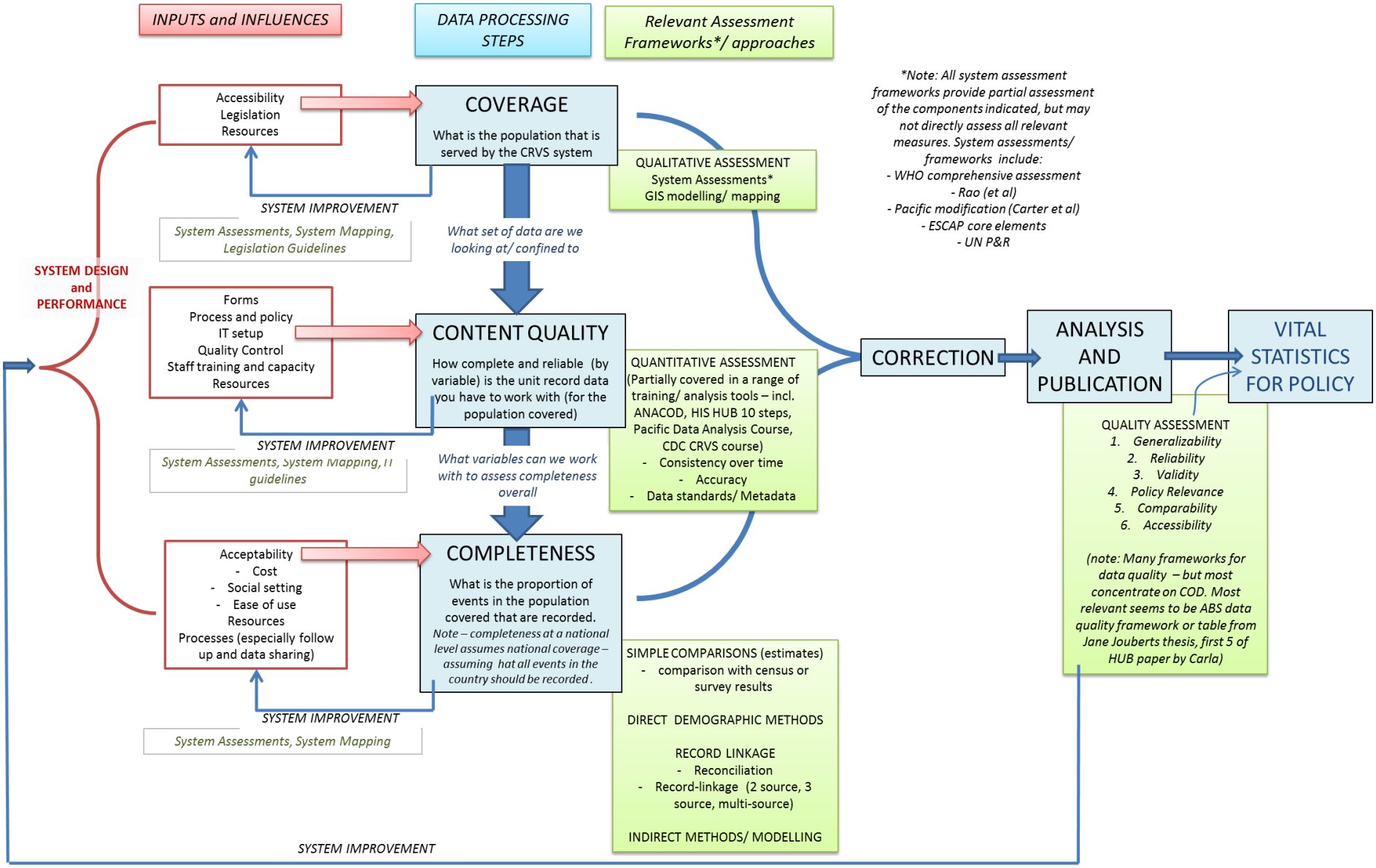


Key concepts – “content completeness”

- Completeness is also frequently used in relation to the individual data fields collected for each record. In other words the proportion of registration records for which specific fields of interest, such as age, sex or cause of death, were properly completed.
- This “content completeness” as it has been called, is a key overall indication of data quality, and directly affects how we can use the CRVS data for analysis; but is also important from the impact it has on our ability to link data records from the register with different sources – which is central to one of the key approaches to assessing completeness of records
- Measuring content completeness can be done either individually by field (such as age or sex), or as a combined assessment of “key fields”.
- In order to complete this calculation there must of course first be an agreed standard on how each field should be filled.

Links between System Design, Data Quality and Vital Statistics for Policy

Pacific Community
Communauté du Pacifique



Measuring coverage

The WHO “Assessment Framework for Improving the Quality and Use of Birth, Death and Cause of Death Information” calculates accessibility as:

- the number of people living in a census enumeration district that have at least one civil registration office or other facility to register vital events, divided by the total national population for the same year, and multiplied by 100 to give a percentage.

However coverage is primarily descriptive…..

Aspect	Considerations
Legislation	What population is covered by the legislation? Is registration compulsory for everyone (i.e. New Zealand), separated by ethnicity of citizenship (i.e. Solomon Islands), or only required in certain circumstances. Does the registration system include events for citizens that occur overseas?
System Design	Is the registry system designed to capture all events or only those in certain locations (i.e. births in the health facility)? Are there separate systems in place for events in hospital versus events in the community?
Geographical and Physical location	How far do people have to travel to complete the registration process? Are access issues compounded by processes that require people to attend in person multiple times to complete the registry process? How accessible is the physical location of the office.
Cost and administrative barriers	Does the overall cost of registering an event preclude access to registration for some people? This should factor in both the absolute cost of registration along with costs associated with travel and obtaining other necessary documents (such as a notice of birth or a statutory declaration). In particular do late registration fees or procedures disadvantage some groups more than others from access.
Infrastructure and Human resources	Do limitations in the operating capacity of some registration offices, such as restricted operating hours, lack of equipment or required forms etc, effectively limit the functionality of some registry centres, thereby affecting coverage in real terms.
Social and cultural factors	Is registration socially acceptable across the community or are there some sections of the population whom do not participate in the process (for example, prior to concerted outreach campaigns in the last several years, there were a number of island communities in Vanuatu that simply did not engage with the local government or registry despite their close physical proximity)? Further, are there social stigmas or processes that discourage registration of events by teenage mothers or single mothers for example?

Adapted from WHO Assessment Framework for Improving the Quality and Use of Birth, Death, and Cause-of-Death Information, the United Nations Principles and Recommendations for Vital Statistics Systems, Rao et. al, and Carter et.al.

Measuring coverage

GIS mapping

- There has been limited documentation of the use of GIS mapping of registry locations, however there is significant potential to use these systems to better understand both system coverage and performance.
- By starting with the locations of registry offices (mappable by type and services provided) and overlaying population data, it should be possible to measure indicators such as:
  - The number of people or proportion of the population within a certain absolute distance of a registry facility; and
  - The number of people or proportion of the population within a certain distance (by road or other specified means of transport) from a registry facility

Our measure of coverage is therefore our best description of the population which we are targeting. This is likely to be couched in terms such as “national” or “sub-national”; “facility” or “community-based”; “inclusive” or “targeted” etc.

Measuring completeness

- At its simplest, measuring completeness is simply the proportion of registered events compared with the “truth”.
- Each of the methods subsequently outlined provide a different approaches to estimating the “true” number of events against which our registered events are then compared.
- The approaches to measuring completeness fall into four primary categories –
  - Direct comparison with a “gold standard” – a source that is considered to be “true”
  - Direct survey of registration completeness
  - Comparison with an expected distribution of events (deaths) by age
  - Comparison with other sources which may also not be complete.

Measuring completeness

- Measuring completeness is done for two distinct reasons
  - Evaluation of the overall performance of the CRVS system / reporting for indicators
  - Evaluation of the vital statistics data for analysis and potential correction
- It is arguable that the level of detail required may differ according to why the assessment is being conducted

- If a system is designed that it only covers events that occur in health facilities, it may be more useful to measure the completeness of the facility based reporting than to try to examine the completeness of all events in the total population in order to examine how well the registry system is functioning.

Direct comparison with a standard

- The simplest measure of completeness
- Direct comparison of the registered events with a second independent source such as national census or use UN estimates from the Demographic Yearbook
- Commonly used for assessing systems that do not have full national coverage such as hospital births or deaths where the Health Information System or ward records may provide an alternative source of data.
- At its simplest, where the census has estimated an actual number of events (births or deaths), the calculation is simply the registered events divided by the reported events in the independent source for the same period multiplied by 100 to derive a percentage.

Direct comparison with a standard

- Where we only have a crude birth or death rate, the number of events must be calculated from this using either the census population or again, a UN estimate.
- Completeness is therefore estimated as:
- YB = RB/(CBR x P) x 100

YB: Estimated registration completeness (%)

RB: Actual number of registered births

CBR: Crude birth rate (per 1000)

P: Total population size (in '000s)

Potential problems:

- The time period for which comparison data is available frequently will not align directly with the time period for which the registry data is being assessed.
  - In this case it may be necessary to apply the crude rate (either birth or death as appropriate) from an earlier time period to a more recent population estimate. This assumes that the rates have remained relatively stable over that time period which may or may not be appropriate,
- The approach assumes that the comparison data itself is complete.

Direct survey of registration completeness

- Common approach for measuring the completeness of birth registration has been through the inclusion of a direct question on surveys such as the Demographic and Health Surveys (DHS)
- The survey asks both "whether the childs' birth was registered" and if the respondent is able to show the interviewer a copy of the certificate.
- Considerations include
  - general survey design and sample frame
  - potential response biases
  - wording of the question and training of interviewers in countries where the "notification of birth" issued by the health service may be mistaken for a formal birth registration
- A modification of this approach is the comparison of the number of students enrolling in school who have a birth certificate compared to the total number of children enrolling in school. Provides a rough estimate only.

Comparison with an expected distribution of events (deaths) by age

- For deaths it is possible to use the age distribution of the registered deaths to assess how complete the data is likely to be.
- This is possible as deaths follow a predominantly standard distribution across age groups, although the methods are not reliable for ages under 5 years.
- Rather than a single approach, this is a family of approaches that vary in the assumptions they make and the extent of data required.
- INDIRECT DEMOGRAPHIC METHODS

Brass Growth Balance

Bennett–Horiuchi

Preston-Coale

- Brass “Growth Balance Method” assumes population stability, so that growth rate is constant with age.
  - Partial Birth rate = growth (r) + 1/C * (partial death rate).

Comparison with an expected distribution of events (deaths) by age

Considerations –

Advantages

- External/ comparison source not required – less resource intensive

Disadvantages

- Is applied across age groups – so does not allow differential results or correction by age
- Method is sensitive to error in the data which changes the slope and to changes in the assumption of stability
- The selection of age groups to “crop” in order to fit the line may also have a significant impact on the estimate of completeness .
  - Murray et.al. examined optimal age ranges against developed country data in 2009. While they found the optimum cut as 40-70, the results from the age trims were found to be highly susceptible to systematic age misreporting, and while not specifically noted in the paper, this would also imply susceptibility to differentials in reporting completeness by age, a scenario possible in many less developed countries.

Multi-source record-linkage

- Commonly called capture-recapture, dual-source comparison, or record linkage studies,
- Compares the overlap of events between two or more sources to estimate the total or “true” number of events without the assumption that either source is correct; thus can be more robust than other approaches.
- It has the further advantage that this technique can be used to examine completeness by sub-group, allowing for the assessment of differentials in reporting completeness by sex, age at death, age of mother and other variables that may be of interest.
- The trade-off is that record linkage can be difficult and as a result this is easily the most resource intensive of the methods described.

Multi-source record-linkage

- The simplest version of this technique, assumes a closed system with independence between the two sources and that all vital events have the same probability of being captured.
- The two source capture-recapture model can be summarised with a two-by-two contingency table where the aim is to estimate the missing value, X, which then allows for a simple estimation of the total population.

Table 4: Two source two-by-two contingency table

		Source 2	
Source 1	Present	Absent	Total
Present	a	b	a + b
Absent	c	X	
Total	a + c		N

$X = bc/a$

$N = a+b+c+X$

(Hook and Regal)

Multi-source record-linkage

- **“Closed” system is unlikely to old absolutely**
- **2 source:** assumes independence which may not be plausible
  - May be possible to use these to set boundaries around what we know regarding the completeness of the data (using reconciled data as a lower limit of the “true “ number of deaths
- **3 sources or more:**
  - LMIC’s are unlikely to have multiple sources at the national level
  - allows for independence but introduces rapidly increasing complexity and choice of models..

- **Towards the next generation of record-linkage studies to advance data quality assessment of civil registration systems in LMIC**
  - ***Washington, D.C., United States, 4-5 April 2016***

Summary of approaches to assessing registration completeness by certainty of results

Increasing certainty of result* ↑

Approach	Application	Method allows for disaggregation by sub-group
Multiple-source record-linkage study	Births and Deaths	Yes
2 source record-linkage	Births and Deaths	Yes
Direct Demographic Measures+	Deaths (for ages >5 years)	No
DHS survey comparison+	Births	No
Direct comparison with a “gold standard”	Births and Deaths	No

A possible approach to selecting a method for measuring completeness at national level?