Get every one in the picture

Data quality and comparison - mortality

Data analysis and Report writing workshop for Civil registration and vital statistics data.

Renee Sorchik
Consultant, Vital Statistics and Civil Registration Expert

UNITED NATIONS ESCAP
Economic and Social Commission for Asia and the Pacific

Why do we need to check the quality of our results?

- Vital statistics are used to monitor and measure the health situation in countries
- Policy makers will use our data to:
  - measure the effectiveness of past policies
  - target interventions and
  - Determine future priorities
- It’s important our data are as accurate and reliable as possible

Checking for plausibility

- A critical step in assessing data quality is determining plausibility – do these results make sense? Are they believable?
- It's important to compare your results to data from:
  - Previous years' vital statistics
  - Census estimates
  - Household surveys (DHS & MICS)
  - Sample registration systems
  - Academic estimates
  - Global Burden of Disease Study

Plausibility of death statistics

- Compare raw counts of deaths and counts by Region, age, and sex to data from previous years and recent census data

- Can you explain any discrepancies?
  - For example, a natural disaster may cause increased number of deaths.

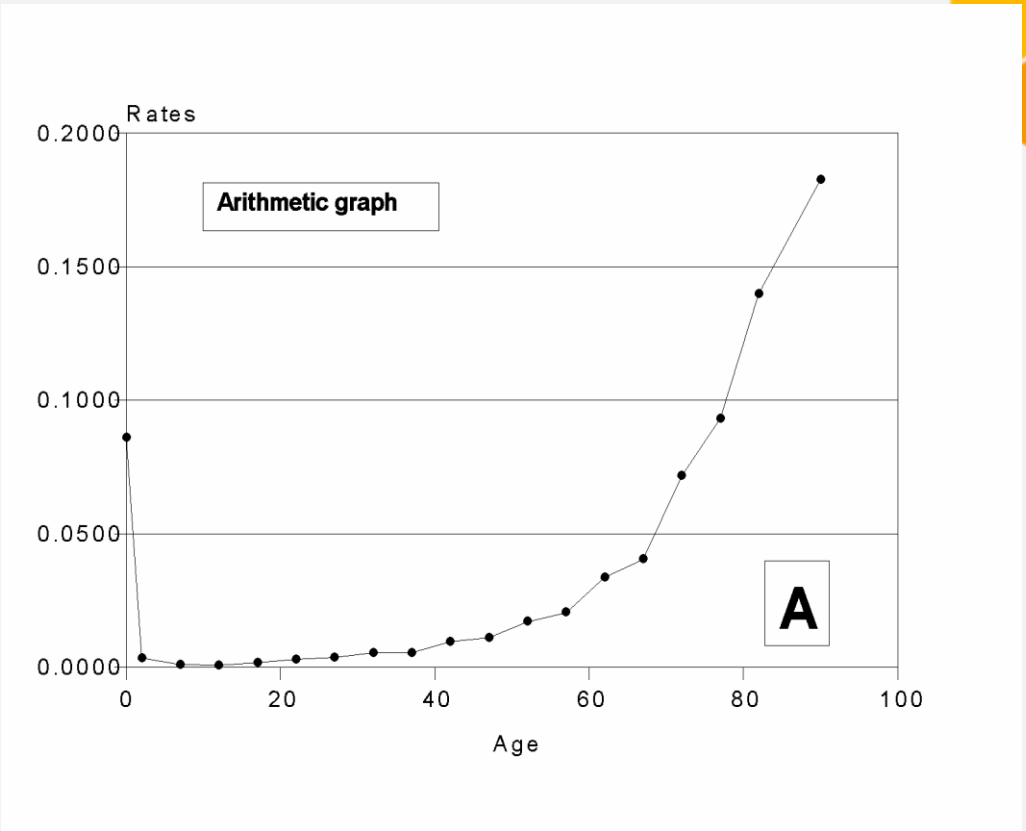
Age-specific death rates are J-shaped

The typical mortality pattern over all age groups has a J-shape.

Mortality is high among infants and young children, after which it declines rapidly, reaching its lower level usually between ages 5-14.

As a general rule,
mortality rates start to increase exponentially beyond age 35 or so.

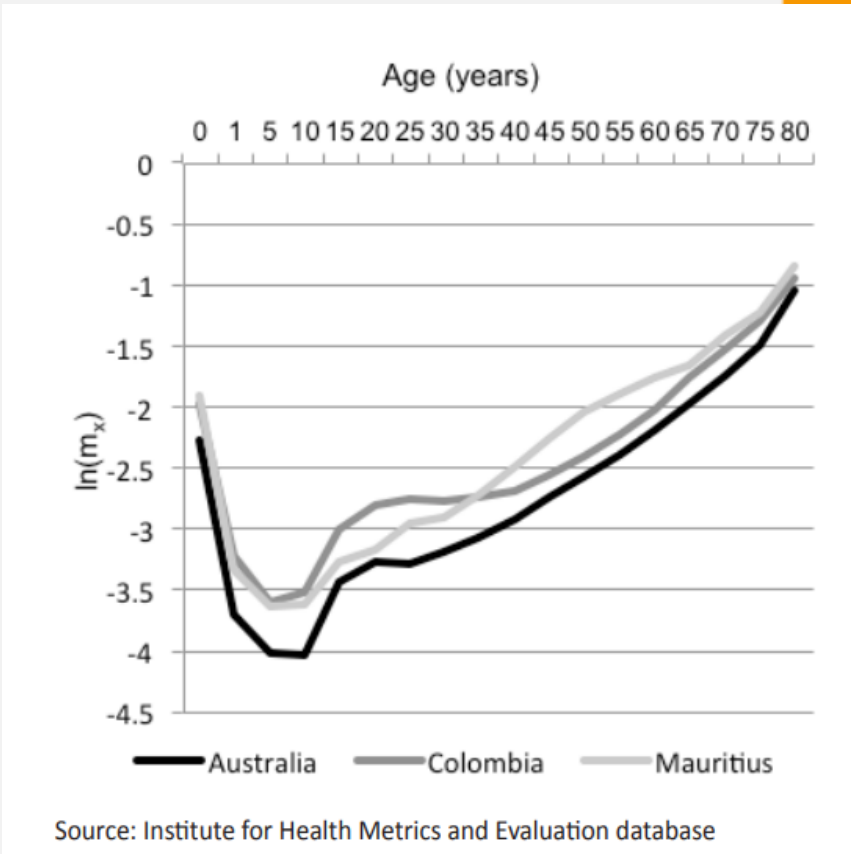
At progressively older adult ages, mortality tends to rise. In this example, it increases rapidly after age 65.



Graph from U.S. Census Bureau's *Population Analysis with Microcomputers Volume I Presentation of Techniques*

Assessing age-specific data quality

- Assess the relative age pattern of mortality among different age groups. Is there anything unexplainable happening on your graph?
- Remember mortality rates are:
- Relatively high around age 0
- Fall and remain low at ages 5-14
- May increase significantly in early adulthood around ages 15-24 due to accidents and incidental causes (and in some cases maternal mortality for women)
- May decrease or remain similar around ages 25-34
- Increases linearly generally from age 35

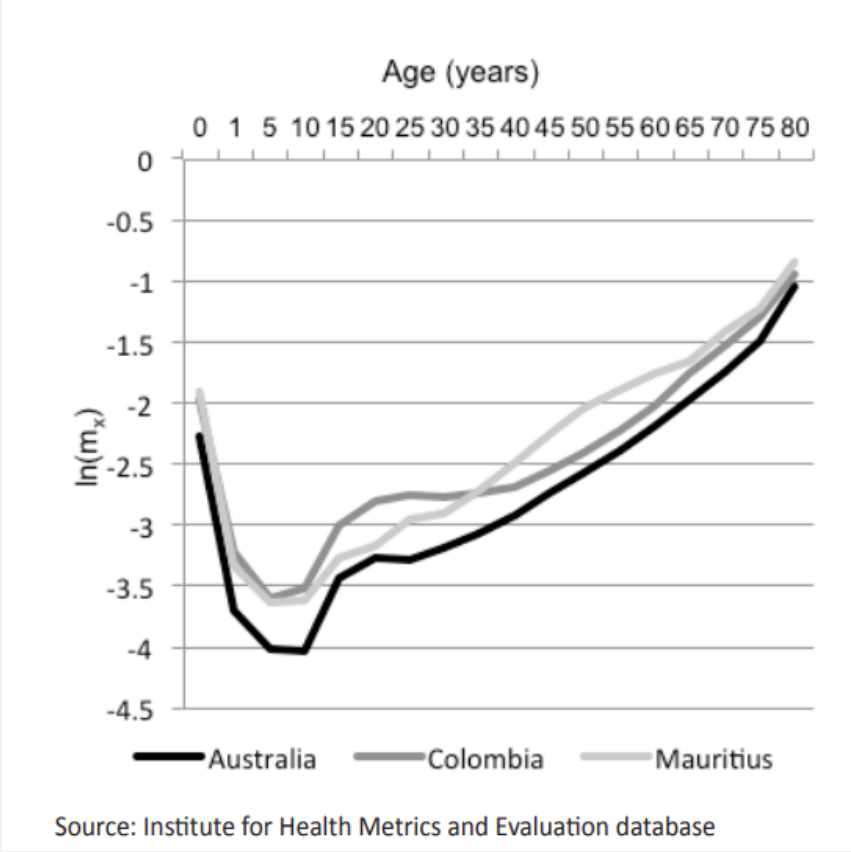


Source: Institute for Health Metrics and Evaluation database

Graph from: University of Queensland Health Information Systems Knowledge Hub's *Mortality statistics: a tool to improve understanding and quality*

Assessing age-specific data quality

- Generally, beyond age 35, death rates rise exponentially which translates to a straight line on logarithmic graphs.
- Departure from this linear pattern suggests deaths are being underreported in certain age groups or that age is being misreported
- Australia —mortality increases smoothly in a straight line from around age 35, as would be expected with more complete death registration.
- Mauritius —Mortality does not increase linearly, from around age 65 the line is not smooth, suggesting underreporting of deaths, particularly at the oldest ages.



Source: Institute for Health Metrics and Evaluation database

Graph from: University of Queensland Health Information Systems Knowledge Hub's *Mortality statistics: a tool to improve understanding and quality*

Checking distribution of deaths by age for plausibility

- University of Queensland Health Information Systems Knowledge Hub's *Mortality statistics: a tool to improve understanding and quality* created a sample guide of how deaths are likely to be distributed by sex and age based on IMRs (see pages 14-16)
- Deviations may indicate under-reporting of deaths at certain ages or mis-reporting of age at death

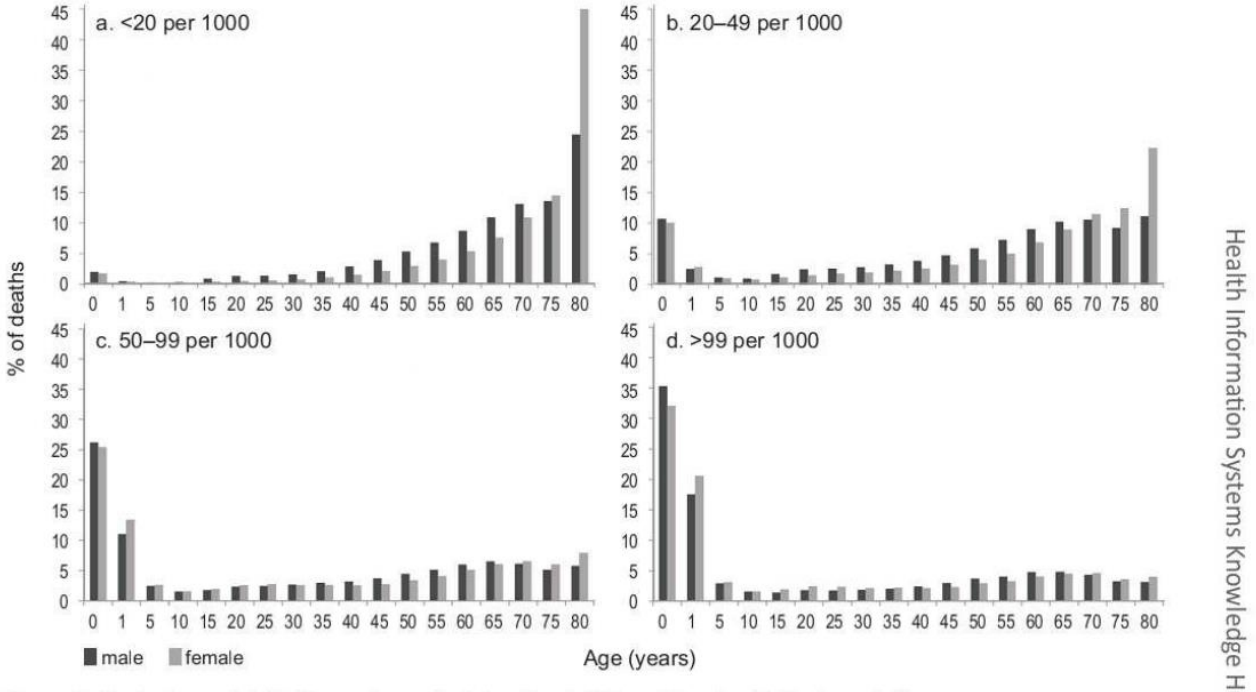


Figure 8 Typical age distributions of reported deaths at different levels of infant mortality

Graph from: University of Queensland Health Information Systems Knowledge Hub's *Mortality statistics: a tool to improve understanding and quality*

Plausibility of NNMR, IMR and U5M values

- Compare to previous years' VS data, data from the census, DHS, MICS, and other sources.
- Graph these values over time
- Graph your data against estimates from the census, DHS, MICS, and any other available sources. How does your data compare?

Are your U5M values plausible?

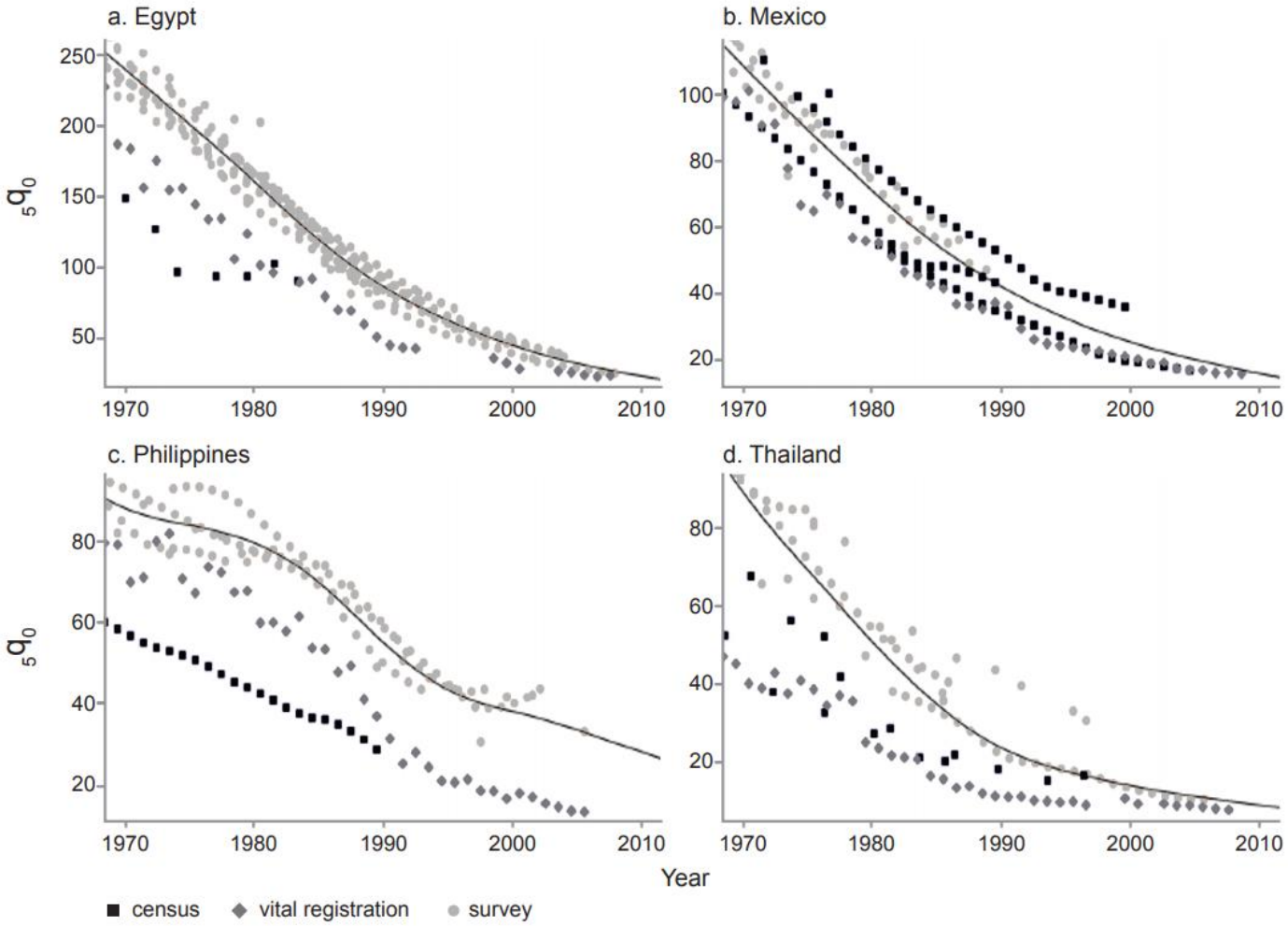
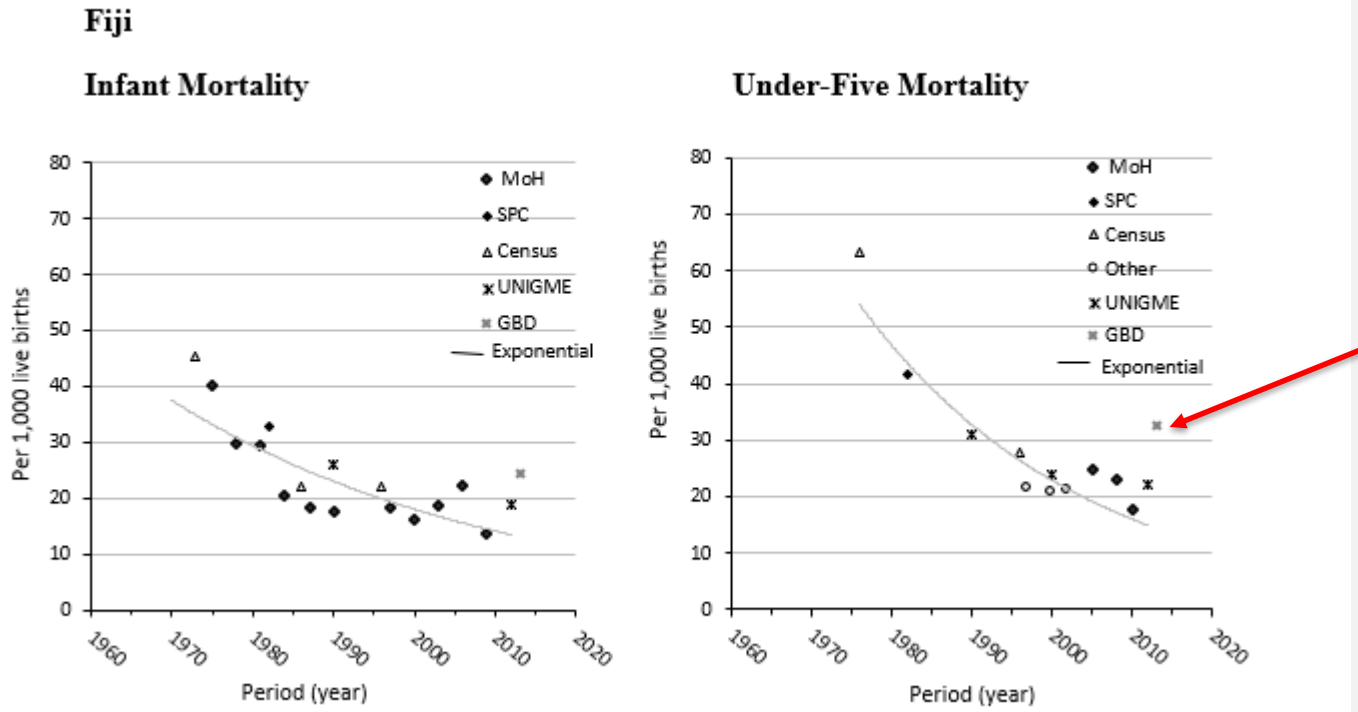


Figure 10 Observed (from vital registration) and estimated levels of the under-five mortality rate, selected countries (1960 - 2000)

Graph from: University of Queensland Health Information Systems Knowledge Hub's *Mortality statistics: a tool to improve understanding and quality*

Are your IMR values plausible?



If this was a VS value, it would require further investigation and explanation.

Graph from: Linhart et al. Mortality Trends in Pacific Island States, June 2014

What to do with implausible values?

- Do your research - was there a policy, event, or other cause that could have affected your data?
- If so, highlight in the text the reason why your number may be different from other estimates, but why you believe it's still accurate.
- If the explanation is more likely due to incomplete data, poorer data quality, or age misreporting, you can still publish your results, but be sure to caution the reader in the text about their low value and mention possible causes for this.
  - It's helpful if you publish aberrant results alongside other estimates for comparison so the reader can interpret them correctly.
- However, you may need to adjust your results before using them as inputs for other mortality measures, such as calculating a life table

Exercises

- Graph the percent distribution of deaths by age group and sex and compare your results with those from the University of Queensland guide based on your IMR.
  - What do you see? Does your data look plausible?
- Review your age-specific death rate logarithmic graph. How does it look? Do you see anything that may require further investigation?
- Graph the additional years of data in the test data IMR Excel tab against you data for:
  - NNMR
  - IMR
  - U5M
  - How does your data compare? Are your results plausible?
- Repeat these exercises with your country data.
- Find other sources of NNR, IMR and U5M data in your country for the last 20 years. Plot these values with your vital statistics data. What kind of trends do you see?