Working Paper No. 17
24 May 2004

ENGLISH ONLY

STATISTICAL COMMISSION and UN ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)

CONFERENCE OF EUROPEAN STATISTICIANS

WORLD HEALTH ORGANIZATION (WHO)

Joint UNECE/WHO/Eurostat Meeting
on the Measurement of Health Status
(Geneva, 24-26 May 2004)

Session 5 – Invited paper

SUMMARY MEASURES OF POPULATION HEALTH

Submitted by World Health Organization, Geneva [*]

I. INTRODUCTION

1. The regular assessment of population health is a key component of the public policy process. Population-level estimates of mortality, morbidity and health states in countries by age, sex and cause, are useful for numerous public health purposes, ranging from the monitoring of new epidemics to measuring progress in reducing old ones for which disease control programmes are in place.

2. To describe health patterns adequately in national or subnational populations according to age, sex and cause, a vast array of estimates need to be generated. It then becomes a major challenge to interpret the key findings of such a review or to compare levels of population health across population groups unless the data are summarised in some fashion.

3. The simplest and most widely used method for producing population health statistics is to aggregate data on individuals in order to generate statistics such as the proportion of the population (or of a particular population sub-group) suffering from a given health problem or living in a particular health state, or the number of individuals who die from a particular cause during a specified interval. This approach rapidly becomes unwieldy when a number of problems or domains are being monitored and

[*] Paper prepared by Colin Mathers.

.

we want to make comparisons over time, across population groups, or before and after some health intervention.

4. There is thus a need for summary indicators of population health. While such summary indicators have many potential uses, there are three that are particularly important for public health policy: (1) to provide easily absorbed summary pictures to convey overall patterns and trends of health to policy makers who do not have the time or interest to absorb and synthesise detailed arrays of information (2) comparisons of the average health levels in different populations or sub-groups, or in the same population over time; and (3) assessments of the relative contributions of different diseases, injuries and risk factors to overall population health. Such summary indicators do not replace the more detailed reporting of data on specific aspects of health and mortality or on the specific causes of health problems; rather they supplement these data with indicators that can be used to monitor trends and compare levels of health across populations.

5. Why consider summary measures of population health at this meeting? Decisions concerning the preferred types of summary indicators for reporting on population health will almost certainly influence not only the reporting on health status data in official statistics, but also the types of information required to be collected and the methods used for collection.

When we ask for summary indicators of health status, what do we want?

6. The demographic and epidemiological traditions take a population perspective as their starting point for the measurement of health status. Many of the challenges in constructing summary indicators of population health are intimately related to the linkage between population and individual health measures. Distinctions between incidence and prevalence perspectives, or period and cohort perspectives, for example, can be recast in terms of different choices as to the set of individuals (real or hypothetical) whose health is aggregated into a population measure. Recent efforts have been made to develop formal expressions of population health as aggregations of individual health measures (1). A helpful starting point is to consider how we would answer the question "Is John healthier than Jack?"

7. Imagine a casual conversation in which one participant says that John is healthier than Jack. What is the common-sense meaning of this statement? How does the use of the phrase ‘is healthier than’ correspond to various measures of individual health? There are at least three distinct and more precise interpretations of this question:

1. Taking into account only *current* levels in various domains of health, is person A in a better state of health than person B?
2. After both person A and person B have died, will person A have lived a healthier life overall than person B?
3. For the remainder of their lives, will person A have a healthier life than person B?

8. Murray, Salomon and Mathers (1) argued that the last question may be closest in meaning to the common usage of the phrase ‘is healthier than.’ Let me illustrate the

distinction between these three perspectives with a simple example. Carol and Patricia are both 35 years old, and in all domains of health (*e.g.*, pain, mobility, cognition, affect, dexterity, vision, *etc*.), they have identical levels at present. Carol and Patricia have had different past health experiences: Carol was paralyzed from the waist down between age 5 and age 34, at which time she underwent successful surgery that resolved her paralysis; Patricia has had full mobility all of her life. Carol and Patricia also differ in terms of current risk factors for health: Patricia has a strong family history of early onset ischaemic heart disease and smokes 2 packs of cigarettes per day, while Carol does not share either risk factor.

9. From the perspective of question 1, Carol and Patricia are in the same current state of health. We do not have sufficient information to answer question 2, but Carol's experience of paralysis from ages 5 to 34 might result in Carol having a less healthy life overall than Patricia, evaluated over their entire lifespans. This depends critically on whether a family history of ischemic heart disease and tobacco consumption will eventuate in a shorter life for Patricia than Carol, and how this shortened longevity would compare to the reduction in health attributable to paralysis. In terms of question 3, we would conclude that Carol is healthier than Patricia because of the known risks to Patricia's health that Carol does not have, all else being equal.

From individual to population level health indicators

10. The discussion of individual level health indicators provides a basis for constructing aggregate health indicators for populations. Taking into account only the current health state of the individual, the first type of population indicator would relate to averages (or indicators of distribution) of the health state indicators. These could comprise vectors of average prevalences for health states by age and sex, or aggregated in some appropriate way across age and sex groups.

11. One potential set of summary indicators for reporting on current health status is a set of prevalence distributions for the core domains of health. An alternative is to choose a single cutpoint on the scale for each domain and give a single prevalence (of incapacity in that domain above a certain level). Such an approach throws away considerable information unnecessarily.

12. A second alternative is to apply a scoring function (health state valuations or scoring function) to provide a severity-weighted prevalence or average health state valuation across all domains. Such a single number provides a way to compare the point-in-time health status of two populations, or of a single population over time.

13. A third alternative is to apply a scoring function (health state valuations or scoring function) to provide a severity-weighted prevalence or average health state valuation across all domains. Such a single number provides a way to compare the point-in-time health status of two populations, or of a single population over time. We discuss these two options further below.

14. The second and third operationalizations of the question "Is A healthier than B?" are much more difficult to measure at population level, since they involve either projections of the future health states of individuals through their life span, or

combinations of past and future health states. In practice, what is used to approximate such indicators are synthetic period indicators, such as life expectancy or healthy life expectancy, constructed in an imaginary cohort which at each age experiences the health state and mortality conditions observed for that age group in a particular period for the population (2, 3).

15. Thus for the second and third types of population health indicator, a form of period health expectancy could be the indicator of choice. Such summary measures of population health (SMPH) are measures that combine information on mortality and non-fatal health outcomes to represent population health in a single number. While such summary measures have many potential uses, there are two that are particularly important for public health policy: (1) comparisons of the average health levels in different populations or sub-groups, or in the same population over time; and (2) assessments of the relative contributions of different diseases, injuries and risk factors to overall population health. These two key uses may best be served by different forms of SMPH (4). This is discussed further in section 4.

16. I should emphasise that the focus in this paper on reporting on population health states and health status (combining health states with mortality risks) does not mean we should not also be interested in the questions "Why is health of A better or worse than B?" and "To what extent do the health states of A and B contribute to their well-being or quality of life?". The answers to these questions require the collection and analysis of information (or indicators) for determinants of health and mortality risk, for disease and injury causes of health and mortality, for health interventions, etc. The need to assess the causal links between such indicators and indicators of health states or health status makes it all the more important to have a conceptually clear distinction between such indicators and indicators of population health state/status.

II. HEALTH STATE PROFILES – INDICATORS BASED ON DOMAIN SCORES

17. Given a survey sample of individual responses about difficulty in a set of health domains, what do we want to know, and thus how should we report on population health? A survey sample of individual responses about difficulty/capacity in a set of health domains, gives us a set of vectors for individuals: (r_1, r_2, r_3, ……., r_n). How do we summarize these into meaningful indicators?

18. There are three alternative approaches to construction of such summary measures:

- Measures that provide a multi-domain profile of individual/population health, e.g. SF-36, HUI, WHS
- Measures that provide an index/weight for preference/value with respect to a given health state, e.g. EQ-6D, HUI, WHS
- Measures that provide a summary measure for population health (SMPH), e.g. HALE, DALY

19. We discuss the first approach in this section - construction of health state profiles as an array of numbers based on domain-specific scores. We use the MOS Short Form 36 (SF-36) to illustrate this approach.

20. One possible approach, appealing in its conceptual simplicity is simply to report prevalences of health states defined by the multiple dimensions and their possible categories defined by responses. However, the potential number of such health states is enormous, even for instruments such as the EQ-5D which many think has inadequate numbers of domain responses (3 per domain). Thus the EQ-5D has 3 to the power 5 categories or 243 categories. Instruments with finer response categories have much greater numbers of potential health states (WHO WHS health module 390,625; HUI Mark III: 972,000; SF-36: over 1 million).

21. Available options for reporting health states measured using the SF-36 include:

- Eight-scale profile
- Norm-based scores (8 scales)
- Summary scales (physical, mental)
- Single summary index (using a preference or valuation function)

22. The eight-scale profile is the most commonly used and recommended approach to reporting on SF-36 based health states for population groups. Figure 1 shows a typical example. The scores for each of the domains are created by summing 0s and 1s for responses to individual questions comprising each domain and then rescaling the score to range from 0 (worst possible responses) to 100 (best possible responses to all questions). Norm based scores may be created from these by rescaling the domain scores for a normative population to have a mean of 50 and a standard deviation of 10.

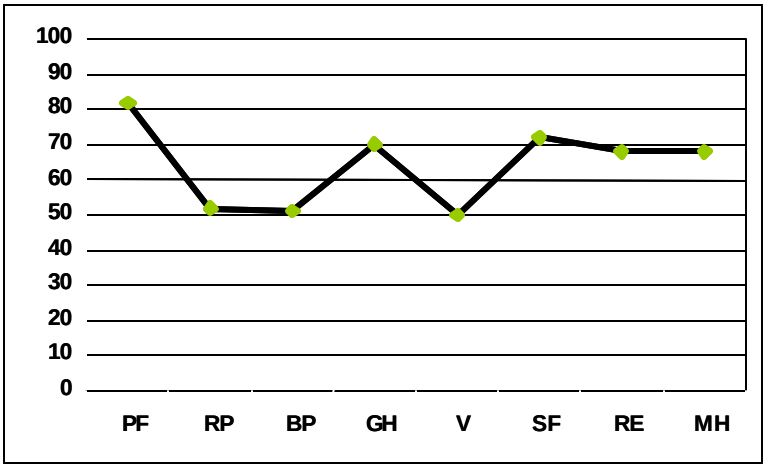

Figure 1. Example of SF-36 health state profile

23. Figure 1 illustrates a number of problems with this approach to summarizing health states. Firstly, although the scores for the domains are plotted on the same vertical axis, the scores across domains cannot be compared with each other. Apart from the arbitrariness of the 0 and 100 setting for each domain, the scales are ordinal not interval and a 10 point increment does not mean the same thing across domains.

24. In comparing two health state profiles, the differences between domain scores for the two groups *describe* change, but are changes on different scales commensurate? How do we conclude whether one group is healthier than another if not all changes are in the same direction across the domains? The solution to this problem requires that we quantify the overall *desirability/preference/levels* of health states.

III. SUMMARY INDEX FOR HEALTH STATE – PREFERENCE OR VALUATION

25. Numerical weights assigned to health states may be 'preferences' derived from various techniques: Visual analog scale (VAS), time trade off (TTO), standard gamble (SG), person trade off (PTO). Alternately, they may be statistical summary scores constructed using techniques such as factor analysis to summarize variation in individual data. A third alternative is arbitrary scoring such as that often used to add up ADL responses. Only the first of these techniques can produce an interval or cardinal-level summary index(5).

26. Ideally, health state valuations, in relation to multidimensional profiles, constitute scalar index values for the overall levels of health associated with different states, measured on a cardinal scale that ranges from zero (for a state equivalent to death) to unity (for a state of ideal health). These valuations formalize the intuitive notions that health levels lie on a continuum and that we may characterize an individual as being more or less healthy than another at a particular moment in time. Health state valuations quantify departures from perfect health, *i.e.*, the reductions in health associated with particular health states. It is important to emphasize that these weights *do not* measure the quality of life of people with disabilities and *do not* measure the value of different people to society.

Health state valuations are discussed further in a following paper in this session.

IV. SUMMARY MEASURES OF POPULATION HEALTH

27. SMPH may provide a convenient and useful summary of the vast array of components of population health. They require the use of health state valuations to summarize the multidimensional health profile of each individual into a single number summarizing the current health status of the individual. SMPH do not replace the more detailed reporting of data on specific aspects of health and mortality or on the specific causes of health problems; rather they supplement and give focus to these data with more comprehensive indicators that can much more easily be used to monitor trends and compare levels of health across populations.

28. Summary measures of population health can be constructed for average levels of population health (such as the WHO healthy life expectancy) and also for distributions of health status within populations. Health equity indicators can be constructed to summarize the distribution of individual health status (somewhat analogous to the Gini index for income distribution) or to summarize differentials in health across strata defined by socioeconomic factors such as income or education.

Health expectancies

29. As summary measures of the overall level of health of a population combining information on mortality and morbidity, health expectancies have two advantages over other summary measures. The first is that it is relatively easy to explain the concept of an equivalent "healthy" life expectancy to a non-technical audience. The second is that health expectancies are measured in units (expected years of life) that are meaningful to and within the common experience of non-technical audiences (unlike other indicators such as mortality rates or incidence rates).

30. We can categorise health expectancies into two main classes: those that use dichotomous health state weights and those that use health state valuations for an exhaustive set of health states. Disability-free life expectancy (DFLE) is an example of the first class (6). Healthy life expectancy (or HALE) is an example of the second class (3). Health state expectancies such as DFLE give an implicit value of zero (equivalent to the valuation of death) for disability above a certain threshold; below this threshold the valuation is 1. This means that the summary indicator is not sensitive to changes in the severity distribution of disability within a population. Additionally, the overall DFLE value for a population is largely determined by the prevalence of the milder levels of disability and comparability between populations or over time is highly sensitive to the performance of the disability instrument in classifying people around the threshold. For these reasons, health state expectancies are not appropriate for use as SMPH (7).

31. In contrast, healthy life expectancy (HALE) is sensitive to changes in the distribution of health states, and thus meets one of the key criteria outlined in (7) for an acceptable SMPH. The simplest method for the calculation of healthy life expectancies is to base them on the currently observed period information: particularly mortality rates and health state prevalences. This allows us to compute health expectancies for populations using Sullivan's method (2-3). This approach requires only the data required for the period life table together with population prevalences for health states which can be measured using cross sectional population surveys and/or burden of disease analyses for prevalent disability in populations.

Health gaps

32. The second main use of summary measures of population health is to assess the relative magnitude of the contribution of different diseases, injuries and risk factors to levels of population health. Such information is a useful input to debates on priorities for research and development, priorities for focused attention of government for policy formulation and for identifying which interventions should be further evaluated. When using summary measures to look at different causes of population health levels, the critical question is different: Is cause A or B a more important contributor to levels of population health. To be comprehensible for the broader public, a summary measure used for such causal attribution should fulfill two important requirements: it should be easily understood and it should have the property of additive decomposition. In other words, the summary measure should be partitionable into causes or sub-groups such that the sum of the SMPH across a set of mutually exclusive and exhaustive categories equals the total.

33. In general, health gaps can be decomposed into the contribution of various causes in a more intuitive and easily communicated fashion than health expectancies. Health gaps are summary measures of population health that measure the gap between current health and some normative standard (8). The best known example of a mortality gap is the PYLL (potential years of life lost) indicator and the best known example of a health gap is the DALY (disability-adjusted life year). DALYs are additive across causes to give the total health gap for a population. A health gap measure such as the DALY thus fulfills different needs for SMPH to be used for causal attribution.

Discussion

34. Wolfson (9) has outlined a vision of a coherent and integrated statistical framework, with a summary measure of population health status at the apex of a hierarchy of related measures, rather than a piecemeal set of unconnected measures. The macro measures at the apex of the system, such as health-adjusted life expectancies, would provide a broad population-based overview of trends and patterns. At the next level, cause-specific summary measures of population health would be used for quantifying the causes of health losses, for identifying the potential for health gain and for linking health interventions to changes in population health. Underlying such SMPH would be summary indicators of population health states and component domain scores for such health states.

35. Such a system should include the capability to 'drill down' below the summary measure to component parts such as incidence rates, prevalence rates, severity distributions, case fatality rates, etc and to health determinants and more distal socio-environmental causes. It should also allow us to 'drill down' below whole of population level to examine inequalities in health and to estimate the impacts of a given intervention on various sub-groups.

36. To achieve such a vision of an integrated system of health indicators, including summary indicators of health states and health status, we must reach a broad consensus on an approach to domain measurement and reporting of such measures, and at least partial consensus on the role of valuation in the construction of the summary indicators.

37. I have argued in this paper that:

- Collecting population data on domains of health and health valuations is feasibleand desirable in developed countries
- Health state profiles (vectors) must be summarized to allow meaningful comparisons of health
- A summary health state index should be preference based not an arbitrary scoring algorithm
- Summary health state index can also be used to compute SMPH that combine health state and mortality data (health status)
- Use of a global valuation function is relevant for comparative purposes

REFERENCES

1. Murray CJL, Salomon JA, Mathers CD. The individual basis for summary measures of population health. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD, eds. Summary measures of population health: concepts, ethics, measurement and applications. Geneva: World Health Organization; 2002.

2. Mathers CD. Health expectancies: an overview and critical appraisal. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD, eds. Summary measures of population health: concepts, ethics, measurement and applications. Geneva: World Health Organization; 2002. (http://www.who.int/pub/smph/en/index.html).

3. Mathers CD, Salomon JA, Murray CJL, Lopez AD (2003). Alternative summary measures of average population health. In: Murray CJL, Evans D, editors. *Health systems performance assessment: debate, new methods, new empricism, and future directions*. Geneva: World Health Organisation, 2003. (http://www.who.int/health-systems-performance).

4. Murray CJL, Salomon JA, Mathers CD, Lopez AD, eds. Summary measures of population health: concepts, ethics, measurement and applications. Geneva: World Health Organization; 2002. (http://www.who.int/pub/smph/en/index.html).

5. Salomon JA, Murray CJL, Ustun TB, Chatterji S (2003). Health state valuations in summary measures of population health. In: Murray CJL, Evans D, editors. *Health systems performance assessment: debate, new methods, new empricism, and future directions*. Geneva: World Health Organisation, 2003. (http://www.who.int/health-systems-performance).

6. Robine JM, Romieu I, Cambois E. Health expectancy indicators. Bulletin of the World Health Organization 1999; 77(2):181-185.

7. Murray CJL, Salomon JA, Mathers CD. A critical examination of summary measures of population health. Bulletin of the World Health Organization 2000; 78(8):981-94.

8. Murray CJL, Mathers CD, Salomon JA, Lopez AD (2002). Health gaps: an overview and critical appraisal. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD, eds (2002). *Summary measures of population health: concepts, ethics, measurement and applications*. Geneva, World Health Organization. (http://www.who.int/pub/smph/en/index.html)

9. Wolfson MC (1998). Measuring health—visions and practicalities. Joint ECE-WHO Meeting on Health Statistics. Rome.

*****