Working Paper No.2
26 October 2005

ENGLISH ONLY

STATISTICAL COMMISSION and UN ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)

WORLD HEALTH ORGANIZATION (WHO)

Joint UNECE/WHO/Eurostat Meeting
on the Measurement of Health Status
(Budapest, Hungary, 14-16 November 2005)

Session 2 – Invited Paper

CRITERIA FOR AND SELECTION OF DOMAINS FOR THE MEASUREMENT OF HEALTH STATUS*

Submitted by the Task Force on Health Status

1. The Task Force on Health Status has prepared criteria for *Developing Criteria for the Selection of Domains*.

2. In considering the domains to be used to define health status, the agreements from the May 2004 meeting should be recalled. These include the adoption of a multi-dimensional approach to health status measurement, and that the focus of measurement is on the health status of individuals and on individuals' capacity to function in each domain.

3. In order to identify the domains of health to be included in the measurement process, a set of criteria for selecting domains first needs to be established. These criteria will help ensure that domains are selected in a standardized, transparent and verifiable manner.

4. Potential criteria were suggested at the May 2004 meeting, and this list has evolved as the work of the taskforce has proceeded. What follows is the current draft list of criteria, grouped under three headings according to whether they relate to issues of relevance, feasibility, or measurement.

5. This document also provides an explanation and rationale for each of the proposed criteria to aid in the selection of the final list.

---

* This paper is Part 2 of 3 papers prepared by the Task Force on Health Status for this meeting. The Task Force consists of representatives from the following countries and international organizations: Australia, Belgium, Canada, Estonia, Hungary, Italy, Netherlands, Norway, Spain, United Kingdom, USA, Eurostat, UNECE, and WHO.

1. Relevance

Face Validity

6. Face validity, in this case, means that the domains of health selected will be immediately seen as plausible and reasonable by ordinary individuals. For example, they should be likely responses to a question such as "when you think of being healthy / ill, what more specifically do you think of?" A fundamental reason for this criterion is that the health status measures need to be credible with a broad public.

7. One corollary of face validity defined in this way is that the domains must be expressible in plain language, using non-technical terms. In turn, this means that the domains are not bio-medically defined diseases, as in the International Classification of Diseases (ICD).

Breadth of domains

8. The domains should span the main aspects of health experienced by the population.This does not mean that they should cover everything (see "parsimony" below); but that they should touch on a wide range of the major aspects of health.

Importance for population health monitoring

9. The domains should all be seen as significant aspects of individuals' health. As such, they should, in combination, be useful for monitoring patterns and trends in population health. Patterns of health include inequalities and differences between population sub-groups at a point in time, while trends refer to changes over time. The domains in combination should be sufficient to support observations as to whether one population is healthier than another, and whether a given population's health is increasing or decreasing over time.

Draw on selected key ideas of the ICF

10. To enhance the relevance of the domains, they should draw on these particular key ideas of the ICF. Specifically, this means that the domains should:

- be agnostic with regard to aetiology;
- be clear that the domains are "components of health";
- focus on "functioning and disability" and not "contextual factors"; and
- focus on capacity and not performance.

11. Note that the domains do not need to bear any particular relationship to the remainder of the ICF, including the detailed classification structure of the ICF in its current form.

2. Feasibility

Feasibility for use in health interview surveys

12. The main purpose of the set of health questions defined by the criteria described here is to collect information on population health, and the main vehicle for doing so is health interview surveys. Thus, one criterion is that the domains be such that the implied questions can be

included on surveys with a reasonable expectation of high quality responses from samples of the general public.

13. Part of survey feasibility is already implied by the criterion of face validity above. A further aspect of feasibility is that the domain allow for the construction of short unambiguous questions. This requires that the domains be succinct and clearly defined, to ensure a consistent interpretation by the lay people from whom the data will be collected. The domains should also be amenable to collection over the telephone, by personal interview or through the mail.

Consistent meaning in different social contexts (cross-cultural comparability)

14. This criterion means that those domains, and the questions developed to measure them, are preferred which are more likely to have close to the same meaning in different populations, especially those from different cultures. This criterion includes, but usually requires more than, accurate translation between languages.

15. Assessing cross-cultural comparability is challenging, and the methods for doing so are currently an active area of research. In turn, applying this criterion may be conceptually or methodologically difficult in current practice.

16. The previous two criteria, feasibility and cross-cultural comparability, together are important reasons for the focus on capacity rather than performance, and on activity rather than participation, in the context of the ICF.

Heterogeneity

17. The domains that are selected must have a reasonable degree of heterogeneity within the population. Domains where everyone is at or near the same level of functioning should be excluded, because they will add almost no information. For example, in Canada there is little variation on the domain describing speech on a national health survey, with 99% of the general population reporting full functioning. This domain would therefore be of less use in describing population health than one with a higher degree of variation within the population.

Parsimony of domains

18. The total number of domains should be small. There are several reasons. One is to keep interview length short, so the question set can become widely used. Another is to reduce the volume of work required for question development, testing, validation, and translation.

3. Measurement

Statistical independence

19. Statistical independence in this context means that in most populations of interest, the levels of health on one domain are unlikely to be correlated with levels of health on another. Achieving complete statistical independence is not a feasible goal. For example, some of the domains may be jointly affected by common external factors such as age or a highly prevalent disease (e.g., musculoskeletal) which jointly affects more than one domain (e.g., pain and mobility). Therefore, although some statistical dependence is to be expected, and can be accepted, the goal remains to minimize such dependence, and to focus on a set of domains that provide the most information on the population's health.

Structural independence

20. Structural independence is distinct from statistical independence. It refers to logical or conceptual dependence among domains. Structural independence applies when an individual's level on one domain in no way pre-determines his or her level on the other domain. An example is pain and seeing. Structural independence entails that domains are not redundant. On the other hand, structural dependence, which should be avoided, occurs when an individual's level on one domain can, to a large extent, be explained or is determined by their level on another domain. An example of such dependence is mobility and usual activities. A lower level of mobility typically results in impairment in usual activities; so as a matter of logic, an individual's levels on the two domains will be correlated. Every effort should be made to ensure that domains are structurally independent.

Clear series of levels within each domain

21. Health status in each domain is characterized by one or another level of capacity to function. The descriptors for these levels should not be too numerous, but there should be enough to describe appropriately the domain in question. The levels should be graded in severity in an ordered fashion. The descriptors of these levels should be expressed in plain unambiguous language, and in a way that anticipates and minimizes differences in interpretation across individuals, across languages, and across cultures.

Within, on or near the skin

22. This criterion means that the domain refers to something that is intrinsic to the individual. In other words, it is independent (to the extent possible) of external factors such as the physical or social environment.

23. This criterion generally is implied by / implies two of the choices with regard to ICF concepts – specifically the focus on "functioning and disability" and not "contextual factors", and the focus on capacity and not performance.

24. This criterion also greatly enhances the prospects for another criterion, cross-cultural comparability.

Suitability for preference measurement

25. Preference measurement involves a separate process wherein individuals are asked to express their choices among health states defined in this case as a series of levels on each of several health domains. The preference function thereby derived can then be used to assign a numerical score to any possible combination of levels on each of the set of health domains. This preference function can then be used to aggregate health survey responses into an overall or summary measure of health.

26. This exercise of defining health domains is not intended to provide health state preferences, nor summary measures of any sort. However, this criterion of suitability means that at some later time, the domains selected have not precluded the possibility that preference measures could be derived.

27. It may be noted that the above criteria of face validity, parsimony, clear levels, consistent meaning, and structural independence are key to this criterion.

Identification of Domains

28. In order to select internationally comparable measures of health status, the domains that conceptualize health in an international context must first be identified. In order to identify such domains, a list of criteria has been developed to ensure that domain selection proceeds in a consistent, objective and transparent fashion. These criteria were outlined in paper 1 and have been used as the base for making recommendations in this document.

29. As a starting point each of the domains that were discussed at the joint UNECE/WHO/Eurostat meeting of May 2004 was included in this preliminary list - they are marked with an asterisk (*). Additional domains have been selected based on a literature review and a review of past surveys and classification systems that measure functional health. In almost all cases we employed concepts that were taken from existing validated instruments. The goal is to obtain a measure of health status that covers the spectrum of functioning including all the significant aspects of health, but that can do so with a limited set of attributes (we recommend no more than 12 core domains with the possibility of having additional domains that could be selected on a per country basis).

30. Each of the domains was assessed against the criteria and we have indicated where we felt that the criteria have been adequately met. Check marks (√) indicate that the criterion is met, tildes (~) indicate a criterion is met with some exceptions, and questions marks (?) indicate criterion about which we are unsure. Blanks indicate that the criterion has not been met. We have also indicated our recommendations about inclusion and have summarized our comments based on whether we think domains should be included, excluded or considered for discussion/clarification at the end of the document. Short forms are used to label the criteria in the table but we assume that the reader will be familiar with paper 1 (Developing Criteria for the Selection of Domains) and will refer to it as required for further clarification of the label headings.

31. We have also divided the selection criteria into two types: (1) criteria for inclusion of individual health domains (e.g., validity, heterogeneity, importance) and (2) more global criteria or criteria that need to be considered in relation to all other domains under consideration (e.g., statistical and structural independence, amenability to preference measurement). The reason for this distinction is that the status of domains with regard to certain criteria (e.g., heterogeneity) will not change as a function of the other attributes in the list, while the status of domains with regard to other criteria (e.g., structural independence) will change as certain other attributes are deleted or added to the list. It is therefore worthwhile to make the distinction to emphasize that as the list is whittled down or built up, the status of domains with regard to the "relational" criteria probably needs to be updated or continually reassessed, to maximize the amount of information considered at each step in the upcoming discussion.

Domain	Criteria												Comment	Recommendation
	Global								Relational					
	Validity	Within skin	Importance	Capacity	Feasibility	Cross cultural	Heterogeneity	Clear levels	Independence	Preference	Parsimony	Breadth		
Physical														
1. Physical functioning: mobility - moving about, moderate or vigorous activity*	√	√	√	√	√	~	√	?		√	√	√	Some structural dependence with usual activities	Include
2. Physical functioning: Upper body movement – lifting and carrying objects	√	√	√	√	√	~	√		√	√	√	√	Some dependence with dexterity	Do not include in primary list
3. Dexterity (use fingers to grasp or handle small objects)*	√	√	√	√	√	√	√	√	√	√	√	√	Criteria met	Include
4. Self care (washing, dressing, taking care of and maintaining general appearance)*	√	√	√	~	√	~	√	?		√	√	√	Possible dependence with physical functioning, dexterity	Discuss possibility of including despite dependence
5. Usual activities (work, shopping, school, social)	√		√		√	?	√	√		√	√	√	Some dependence with physical functioning, anxiety, affect	Do not include, indirectly covered in other domains
6. Vitality/fatigue (tiredness, lack of energy or feeling rested and refreshed)*	√	√	√	√	√	√	√	√		√	√	√	Possible dependence with affect and sleep	Include
7. Sleep (difficulty falling asleep, waking up early, or frequently during the night)	√	√	√	√	√	√	√	~	~	√	√	√	Dependence with vitality/energy, affect	Consider not including and measuring sleep's effects with vitality/energy domain
8. Breathing	√	√		√		√		√		√	√	√	Criteria not met	Do not include – breathing can be covered under discomfort
9. Urinary incontinence		√		√	√	√	√	√	√	√		√	Criteria not met	Do not include
Mental														
10. Cognition – memory and concentration*	√	√	√	√	√	√	√	√		√	√	√	Dependence with thinking and problem solving	Consider including one measure of cognition
11. Cognition - thinking, problem solving	√	√	√	√	√	√	√	√		√	√	√	Dependence with memory	Consider including one

Domain	Criteria												Comment	Recommendation
	Global								Relational					
	Validity	Within skin	Importance	Capacity	Feasibility	Cross cultural	Heterogeneity	Clear levels	Independence	Preference	Parsimony	Breadth		
													and concentration	measure of cognition
12. Alertness	√	√		√		√	√	√		√		√	Dependence with energy/vitality, cognition	Do not include
13. Communication	√	√	√	√	√	√	√	√		√	√	√	Dependence with speech and hearing	Do not include, covered with speech and hearing
Emotional														
14. Affect (emotional state- happiness, depression)*	√	√	√		√	?	√	√	√	√	√	√	Criteria partially met, the degree to which this measures capacity is an issue	Include, many of the feeling domains do not deal directly with capacities but are still intrinsic, within the skin
15. Psychological functioning (possibly mastery, coherence, outlook on life)*													From the May meeting but it was not further defined. Potential dependence with anxiety, affect and cognition.	Requires clarification
16. Anxiety (nervousness, uneasiness, worry, concern or fear or stress)*	√	√	√		√	√	√	√		√	√	√	Criteria partially met, the degree to which this measures capacity is an issue; possible perception of overlap with affect and dependence with social relationships	Include, many of the feeling domains do not deal directly with capacities but are still intrinsic, within the skin
17. Self-esteem		√	√	√	√		√	√		√	√	√	Dependence with interpersonal relationships, affect	Consider covering indirectly through interpersonal relationships
18. Relaxation and leisure	√		√	√			√	√		√	√	√	Dependence with anxiety, affect	Do not include
Sensory														
19. Vision (ability to see)*	√	√	√	√	√	√	√	√		√	√	√	Including recognition implies some dependence with cognition	Include as is and accept possible dependence
20. Hearing (ability to hear)*	√	√	√	√	√	√	√	√		√	√	√	Including understanding	Include as is and accept

Domain	Criteria												Comment	Recommendation
	Global								Relational					
	Validity	Within skin	Importance	Capacity	Feasibility	Cross cultural	Heterogeneity	Clear levels	Independence	Preference	Parsimony	Breadth		
													implies some dependence with cognition	possible dependence
21. Speech (capacity to speak)	√	√	√	√	√	√		√		√	√	√	Criteria partially met, dependence with communication	Include
22. Pain and discomfort*	√	√	√		√	?	√	√		√	√	√	Criteria partially met, the degree to which this measures capacity is an issue	Include, many of the feeling domains do not deal directly with capacities but are still intrinsic, within the skin
23. Taste and Eating		√		√		√		√	√	√	√	√	Criteria not met	Do not include
24. Smell		√		√		√		√	√	√	√	√	Criteria not met	Do not include
25. Touch		√		√		√		√	√	√	√	√	Criteria not met	Do not include
Interpersonal														
26. Interpersonal/Social relationships (capacity to sustain relationships with strangers, formal/informal relationships, family)*	√	√	√	√		√	√	√		√	√	√	Criteria partially met, dependence with anxiety	Discuss way to characterize the domain so that dependence is reduced as it is a key domain
27. Social functioning*													A suggestion from the May meeting but it was not further defined	Requires clarification
28. Social Support	√		√		√	√	√	√	√	√	√	√	Not intrinsic to the individual	Do not include
29. Reproductive Functioning	√	√	√	√			√	√	√	√	√	√	Criteria not met, can be captured through social relationships	Do not include
30. Sexual Functioning	√	√	√	√		√	√	√	√	√	√	√	Criteria partially met	Consider for inclusion

Summary of Recommendations

A. Recommended for inclusion:

1. Physical functioning: a) mobility
2. Dexterity
3. Vitality/Fatigue
4. Affect
5. Anxiety
6. Vision
7. Hearing
8. Pain and Discomfort
9. Cognition: a) memory and concentration, b) thinking and problem solving
10. Social relationships and functioning, including aspects of communication

B. Recommended for exclusion from the core list:

1. Physical Functioning: b) upper body movement
2. Self care
3. Sleep
4. Psychological functioning – covered under affect and anxiety
5. Self-esteem - could be incorporated into social relationships
6. Sexual functioning - could be incorporated into social relationships
7. Speech
8. Usual activities
9. Breathing
10. Urinary incontinence
11. Alertness
12. Relaxation and leisure
13. Taste and eating
14. Smell
15. Touch
16. Social support
17. Reproductive functioning
