IX. RESPONSIVENESS: LEVEL AND DISTRIBUTION

1. WHR 2000

Responsiveness was included as an intrinsic goal in the health systems performance framework because the way people are treated when they come into contact with the system can improve or reduce their well-being independently of health outcomes. There were seven domains of responsiveness: dignity, autonomy, confidentiality of information, prompt attention, access to social-support networks, quality of basic amenities, and choice of health-care provider. For WHR 2000, WHO obtained data on responsiveness through key informant surveys for 35 out of 191 countries. Within each of the 35 countries, a single focal person was canvassed. Each focal person selected an average of 50 key informants from a broad range of health-system stakeholders, including consumer groups, to answer a short questionnaire. Focal persons oversaw data capture and submission of data to WHO. Data from two such surveys was not of sufficient quality to be used in WHR 2000.

For the overall measure of responsiveness, key informants were asked to provide a general rating of the health system in their country with respect to the seven domains after they had answered specific questions on each domain. The specific questions were used to ensure that key informants correctly identified the various components of the domains. Correlation and exploratory factor analysis were undertaken to check for consistency. There was a high degree of consistency between the responses to the specific domain questions and the general rating questions.

In the final analysis, the overall domain scores were regressed on a set of covariates for the 33 surveyed countries, and from these regressions the missing data for the remaining 158 countries were imputed. For the distribution of responsiveness, key informants were asked to identify marginalized groups. This information was used together with the information on the size of those groups in the country to develop a responsiveness inequality score (distribution). Once again, the information was imputed for the 158 non-surveyed countries.

2. Main commentaries and criticisms

Data sources

Many criticisms were raised on this method of obtaining information (Almeida et al. 2001; Navarro 2001; Blendon et al. 2001a) and estimating the missing data (Williams 2000; Almeida et al. 2001; Aalto 2000). Criticisms included the fact that the method was biased because most of the key informants were WHO people; that the method was inherently flawed as it was not a representative sample of the population; that only seven questions out of 42 were used for the index; and that too many imputations were made from too little data. In particular, it was noted that the data and methods used to estimate responsiveness inequality for the unsurveyed countries resulted in multiple tied ranks (Williams 2000).

Relative weight of responsiveness

Commentators from regional consultations questioned the relative importance of having responsiveness in the framework (WHO Regional Office for the Eastern Mediterranean 2001).

Domain weights

Several commentators questioned the relative weights of the seven domains in the aggregation for an index of overall responsiveness (WHO Regional Office for Europe 2001; WHO Regional Office for South-East Asia 2001; WHO Regional Office for the Western Pacific 2001; WHO Regional Office for Africa 2001).

Responsiveness of the broader health system

Comments were made in regional consultations (WHO Regional Office for Europe 2001; WHO Regional Office for South-East Asia 2001) that responsiveness needed to reflect the broader boundaries of the health system, including public access to information and other services of health protection and promotion (see also Ugá et al. 2001; Travassos 2001; Oswaldo Cruz Foundation 2000).

Sources of information

The Blendon et al. (2001a) critique addresses the issue of who is better qualified to judge health-care systems – key informants or users of the system. Blendon et al. (2001b) state that both satisfaction and responsiveness measures are important when information is canvassed from the population.

Translation, validity and reliability

The critique of Aalto (2000) covers questions related to translating the concept of responsiveness and cross-cultural validity. The issue of cross-

cultural validation was also raised in several regional consultations (WHO Regional Office for Europe 2001; WHO Regional Office for South-East Asia 2001; WHO Regional Office for the Western Pacific 2001). Participants in the regional consultations (WHO Regional Office for Africa 2001; WHO Regional Office for the Western Pacific 2001) felt that translation might be a slightly more difficult problem for responsiveness than for other modules owing to the abstractness of the concepts involved (see also Almeida et al. 2001). Aalto (2000) and participants in regional consultations (WHO Regional Office for South-East Asia 2001; WHO Regional Office for the Western Pacific 2001) criticized the availability of standard instrument psychometric data on the responsiveness key informant instrument. Aalto (2000) and the SEARO and WPRO regional consultations (WHO Regional Office for South-East Asia 2001; WHO Regional Office for the Western Pacific 2001) indicated that this type of data should be available for any subsequent responsiveness questionnaire instruments (e.g., in household surveys).

Universality of domains

Aalto (2000) commented extensively on the need to provide a convincing rationale for the choice of domains. The change to household surveys was commended but WHO was cautioned that cross-cultural validation of survey questions on domains should be ensured in any future household survey work. (To some extent this is linked to the issue of translation.) At some of the regional consultations, participants raised the issue of relevance of the domains in different cultural settings (WHO Regional Office for Africa 2001; WHO Regional Office for the Western Pacific 2001; WHO Regional Office for South-East Asia 2001). The critique of Williams (2000) also touched on this issue.

Non-users

A commentary of the Brazilian Ministry of Health (Oswaldo Cruz Foundation 2000) criticized the responsiveness work on the grounds that the WHO indicator was limited to measuring the experiences of people who actually use health services.

3. WHO responses and proposals

Data sources

In order to improve data sources, WHO has focused on developing survey instruments to obtain information from households. The number of countries covered by household surveys will be increased substantially. Some 60 countries have already been surveyed through the Multi-Country Survey Study, and the World Health Survey will cover a further 70 countries. Using

the Multi-Country Survey Study data, distributional measures of responsiveness are being developed and tested.

Relative weight of responsiveness

The relative importance of responsiveness within the overall framework is being tackled with new survey questions in the World Health Survey, which are currently being tested.

Domain weights

Since WHR 2000, WHO has launched the Multi-Country Survey Study in which households were asked directly about their relative weights for each domain. In analyzing the data from this Study, WHO has found that they indicate a common set of rankings of domains across countries, and possibly a tendency towards a common set of weights. However, conclusions on the weight structure across countries are limited by the structure of the original question. New questions to elicit weights from respondents for the domains are being tested in the World Health Survey pilots.

Responsiveness of the broader health system

New questions on heath promotion and support structures for families looking after ill family members at home are being tested in the World Health Survey pilots.

Satisfaction versus responsiveness

The WHO responsiveness survey module is designed using the latest thinking in the field of patient assessment measurement, based on patients' interactions with the health-care system. Satisfaction remains an interesting measure for other reasons because it solicits people's opinions about the system, rather than their reports of personal interactions with it. More work is being done to test the use of techniques to improve cross-population comparability of results from surveys of people's experiences.

Translation, validity and reliability

Since WHR 2000, the WHO has developed an extensive translation protocol for the Multi-Country Survey Study, which has been improved further in the piloting of the World Health Survey instrument. In addition, facility surveys are being developed to collect evidence on the validity of the questionnaire instrument. These surveys will enable the comparison of observations on certain domains of responsiveness in facilities with reports from individuals using those facilities. Other standard validity strategies recommended by Aalto (2000) and mentioned by participants in regional consultations (WHO Regional Office for South-East Asia 2001; WHO Regional Office for the

Western Pacific 2001), such as comparisons with similar data series, are being pursued. WHO is also continuing to document the results of standard psychometric tests of the household survey instruments.

Universality of domains

WHO has produced a paper documenting the criteria for selection of the domains (De Silva 2000). Since WHR 2000, an eighth domain of responsiveness – clarity of communication – has been included. Questions relating to this new domain were developed in response to consultations and included in the Multi-Country Survey Study. A group of ethicists was asked to review the cross-cultural dimensions of the responsiveness domains. Their findings were submitted to the technical consultation on responsiveness (World Health Organization 2001). In addition, further work is currently underway to map the responsiveness domains to UN and other international conventions and treaties on human rights. More cognitive testing is planned for the responsiveness module items in the World Health Survey.

Non-users

With respect to this critique, efforts have concentrated on finding ways to include non-user and low-user groups. As a first attempt at addressing the non-user and low-user problem, models to predict responsiveness for non-users and low-users were developed. This proposal was discussed at the technical consultation on responsiveness (World Health Organization 2001), and goes some way to addressing the problem of non-users and low-users. Both the Multi-Country Survey Study and the piloted World Health Survey instruments have included questions regarding utilization.

4. SPRG comments and recommendations

Data sources

SPRG members agree with the criticisms made by external commentators about the data sources. In particular, SPRG members concur that people using health systems should be asked their opinion about it, rather than relying on information from key informants. They recommend that if the indicator of responsiveness is to be utilized in future, it will be necessary for WHO to obtain representative household-level data for all countries.

Relative weight of responsiveness

Some SPRG members were concerned with the inclusion of 'responsiveness' for the evaluation of health-systems performance. Responsiveness as defined by WHO is meant to deal with the interactions of users with the health system, and includes features such as respectful treatment,

confidentiality of information, prompt attention, and involvement in decision-making. Such features apply to many service activities, e.g., educational services, transportation services, etc. Some SPRG members therefore felt that it might be better to deal with issues of responsiveness generically, i.e., at the national (or even international) level rather than at the health-system level. Other members rejected this notion and pointed out that if responsiveness was measured at the national level, accountability could only be attributable at that level – and not at the health-system level.

SPRG members were also surprised that the weight on responsiveness in the composite indicator was as high as 25%, the same weight as for average health level. This implies that a one-point increase in responsiveness is valued as highly as a one-point increase in the scale used for health (equivalent to almost one year of health-adjusted life expectancy). In view of the implied trade-off, SPRG members wondered whether appropriate questions had been asked to elicit the relative weights for responsiveness and the intrinsic goal of average health level. SPRG members also wondered whether it might have been appropriate to incorporate a changing set of weights for responsiveness at different stages of health-system development. (The present set of weights is constant for all levels of development.) It could be argued that a greater weight attaches to pure health goals relative to responsiveness at low levels of life expectancy (e.g., 50 years) than at high levels of life expectancy. Once life expectancy reaches 70 or more years, as in the OECD countries, it may be more appropriate to use a relatively larger weight on responsiveness. At high levels of life expectancy the room for further improvement in health is limited, and other goals – such as responsiveness of the health system – may assume greater importance.

Domain weights

Several commentators (WHO Regional Office for Europe 2001; WHO Regional Office for South-East Asia 2001) suggested that the responsiveness domains should be given country-specific weightings in the aggregation of the domains into an index of overall responsiveness (while maintaining the relative weight of overall responsiveness *vis-à-vis* other goals of the health system). SPRG members also recommend that WHO experiments with a non-linear system of weights to reflect changing priorities that might attach to responsiveness relative to pure health goals at different stages of development. It was recognized that some of the responsiveness measures deal with human rights issues, such as dignity and confidentiality, which need to be addressed at all stages of development. An appropriately specified non-linear system of weights can accommodate constant linear weights on certain domains of responsiveness.

Some SPRG members suggested that WHO assess the relationship between the level of responsiveness (by domain) in a country and the level of financial resources available to its health system. This approach will help assess

whether there is a differential capacity in countries for producing responsiveness.

Responsiveness of the broader health system

Some SPRG members also questioned the use of the term 'responsiveness'. Responsiveness of the health system could be construed to include several other features apart from interactions with the population – such as the delivery of health services, health promotion and protection, and health education. The term has often been confused with the notion of how well the health system 'responds'. In consequence, some SPRG members suggested that WHO should consider changing the term 'responsiveness' to something like 'interactions with users'. Other SPRG members suggested that possibly a term like 'patient-/people-centredness' or 'patient rights' might work.

Some SPRG members suggested that WHO should conduct a thorough survey to identify potential questions that address responsiveness as it relates to health promotion and disease prevention.

Based on the critiques and their own assessment, SPRG felt that the present WHO questionnaire on responsiveness was geared to eliciting information mainly on personal health services, and that health promotion and protection activities were relatively neglected. Some SPRG members also wished to see the responsiveness of financing activities assessed.

Taken together, SPRG felt that there was a case for extending the responsiveness domains to aspects of health-system activities beyond personal health-care services – e.g., early warning systems in the case of epidemics or other biological or environmental health threats, health promotion and protection, health education in schools, research, etc.

Satisfaction versus responsiveness

SPRG agrees with Blendon et al. (2001a) that users rather than key informants should be the judges of the health-care system. SPRG acknowledges the usefulness of satisfaction measures in general, but feels they are not necessarily a substitute for responsiveness in the framework of HSPA. For example, a person might feel satisfied because he was cured, but he may not have received prompt attention or have been treated with respect. Alternatively, a person might feel satisfied if he were prescribed drugs, even if these drugs were unnecessary, or harmful, for his condition.

In this regard it should be noted that, unlike responsiveness, measures of satisfaction do not adjust for people's differing expectations of the health system. This adjustment is made through the HOPIT approach **(**see Section XIII on Cross-Population Comparability).

SPRG recommends that WHO should continue work on developing experience measures and the use of vignettes and other techniques for dealing with cultural differences in expectations and response tendencies.

Translation, validity and reliability

Ensuring the accuracy of translation is a difficulty faced by all surveys administered in multiple languages. In particular, in a country with many dialects, the infeasibility of issuing a questionnaire in all its languages and dialects presents obvious problems in administering interviews. (For example: Was the interviewer able to communicate in the respondent's dialect? How well was the interviewer able to translate concepts and questions on the spot? Did he use exactly the same wording for different households?) SPRG members as well as participants in the regional consultations (WHO Regional Office for Africa 2001; WHO Regional Office for the Western Pacific 2001) felt that translation might be a slightly more difficult problem for responsiveness than for other modules owing to the abstractness of the concepts involved.

SPRG recommends that WHO conducts more extensive cognitive testing to evaluate how respondents interpret the survey items. In addition, rigorous interviewer training protocols need to be developed, tested and applied. Training and management of interviews must meet high standards to try to ensure the consistent application of the survey protocols.

Universality of domains

SPRG recommends that WHO should document the mapping of cultural influences on responsiveness domains as well as the mapping of domains onto UN and other international conventions and treaties on human rights.

Non-users

In noting the Brazilian comments (Oswaldo Cruz Foundation 2000, Travassos 2001), SPRG felt that there were indeed serious problems in using an indicator that was limited to measuring the experiences only of people who use health services, especially when making cross-country comparisons. For example, it could turn out that only 20% of the population of country A used its health-care system, and this system was judged to be perfectly responsive by its users (according to the scoring criteria). In contrast, in country B, 80% of the population used its health-care system, which was judged to be only 50% responsive by its users. Which system is more responsive?

According to the WHR 2000 definition of responsiveness (experience of users), the health-care system of country A is more responsive. Several SPRG members expressed unease with this logical conclusion. However, the conclusion is inevitable if coverage is not a part of the definition of

responsiveness. Indeed, according to some SPRG members, the term 'responsiveness' evokes the idea of a health-care system responding to people's needs. Hence, if people in country A have been put off from using the system (because of out-of-pocket costs, lack of knowledge, high transport costs, previous bad experience, etc.), this should be reflected in any measure of the responsiveness of the system.

SPRG noted the development of an approach by WHO to predict responsiveness among non-users. However, if the WHO maintains its current approach to measuring responsiveness among the actual users of the system, SPRG recommends that measures of responsiveness should be accompanied by measures of utilization.

5. References

Aalto, A. M. (2000): Measuring the responsiveness of health care system in the World Health Report 2000. In eds. Häkkinen, U. and Ollila, E. *The World Health Report 2000: What does it tell us about health systems? Analyses by Finnish Experts*. Helsinki, Finland: National Research and Development Center for Welfare and Health (STAKES). [http://www.stakes.fi/english/publicati/Publications.htm]

Almeida, C., P. Braveman, M. R. Gold, C. L. Szwarcwald, J. M. Ribeiro, A. Miglionico, J. S. Millar, S. Porto, N. R. Costa, V. O. Rubio, M. Segall, B. Starfield, C. Travassos, A. Ugá, J. Valente, and F. Viacava (2001): Methodological concerns and recommendations on policy consequences of the World Health Report 2000. *Lancet,* 357(9269): 1692-1697.

Bernard, R. H. (1994): *Research Methods in Anthropology. Qualitative and Quantitative Approaches*. AltaMira Press.

Blendon, R. J., M. Kim, and J. M. Benson (2001a): The public versus the World Health Organization on health system performance. *Health Affairs,* 20(3): 10-20.

Blendon, R. J., M. Kim, and J. M. Benson (2001b): Authors respond to WHO critics. *Health Affairs,* 20(4): 253.

Darby, C., N. Valentine, and C. J. L. Murray (2000): WHO strategy on measuring responsiveness. Global Programme on Evidence for Health Policy Discussion Papers, No. 23. Geneva, Switzerland: World Health Organization.

De Silva, A. (2000): A framework for measuring responsiveness. Global Programme on Evidence for Health Policy Discussion Papers, No. 32. Geneva, Switzerland: World Health Organization.

De Silva, A. and N. Valentine (2000b): Measuring responsiveness: results of a key informant survey in 35 countries. Global Programme on Evidence for Health Policy Discussion Papers, No. 21. Geneva, Switzerland: World Health Organization.

Häkkinen, U. and E. Ollila, eds. (2000): *The World Health Report 2000: What does it tell us about health systems? Analyses by Finnish experts.* Helsinki: National Research and Development Center for Welfare and Health (STAKES). [http://www.stakes.fi/english/publicati/Publications.htm]

Murray, C. J. L., J. Frenk, D. Evans, K. Kawabata, A. Lopez, and O. Adams (2001): Science or marketing at WHO? A response to Williams. *Health Economics,* 10(4): 277-282.

Murray, C. J. L., K. Kawabata, and N. Valentine (2001): People's experience versus people's expectations. *Health Affairs,* 20(3): 21-24.

Navarro, V. (2002): The World Health Report 2000: Can health care systems be compared using a single measure of performance? *American Journal of Public Health,* 92(1): 31-34.

Navarro, V. (2001): World Health Report 2000: Response to Murray and Frenk. *Lancet,* 357(9269): 1701-1702.

Navarro, V. (2000): Assessment of the World Health Report 2000. *Lancet,* 356(9241): 1598-1601.

Oswaldo Cruz Foundation, Brazilian Ministry of Health (2000): Report of the workshop "Health Systems Performance: The World Health Report 2000". 14-12-2000, Rio de Janeiro.

Travassos, C. (2001): Assessing health systems performance -- a critical appraisal about the WHO World Health Report 2000 and future developments. Paper presented at Conference on Restructuring of Health Services and Corporate Public Health in the Era of Reforms. 5-7-2001, Maastricht, The Netherlands.

Ugá, A. D., C. M. Almeida, C. L. Szwarcwald, C. Travassos, F. Viacava, J. M. Ribeiro, N. R. Costa, P. M. Buss, and S. Porto (2001): Considerations on methodology used in the World Health Organization 2000 Report. *Cadernos de Saude Publica,* 17(3): 705-712.

Ustün B., S. Chatterji, M. Villanueva, L. Bendib, C. Celik, R. Sadana, N. Valentine, C. Mathers, J. P. Ortiz, A. Tandon, J. Salomon, Y. Cao, X. W. Jun, and C. J. L. Murray (2000): WHO multicountry household survey study on health and responsiveness 2000-2001. Global Programme on Evidence for Health Policy Discussion Papers, No. 37. Geneva, Switzerland: World Health Organization.

Ustün, B. and C. J. L. Murray on behalf of the World Health Organization Survey Programme (2001): World Health Survey: objectives, design, modules. Global Programme on Evidence for Health Policy, mimeo. Geneva, Switzerland: World Health Organization.

Valentine, N., J. P. Ortiz, A. Tandon, K. Kawabata, and C. J. L.Murray (2001): Health System Responsiveness: Evidence from Population Surveys in 15 Countries. Global Programme on Evidence for Health Policy, mimeo. Geneva, Switzerland: World Health Organization.

Valentine, N., J. Salomon, and C. J. L. Murray (2001): Weights for responsiveness domains: analysis of country variation in 57 national sample

surveys. Global Programme on Evidence for Health Policy, mimeo. Geneva, Switzerland: World Health Organization.

Williams, A. (2000): Science or marketing at WHO? A Commentary on 'World Health 2000'. *Health Economics,* 10(2): 93-100.

World Health Organization (2001): Report on WHO meeting of experts: Responsiveness. 13-9-2001, Geneva, Switzerland.

WHO Regional Office for Africa (2001): General Report on the Regional Consultative Meeting on Health Systems Performance Assessment: Final Report. 18-5-2001, Harare, Zimbabwe.

WHO Regional Office for the Americas (2001): Regional consultation of the Americas on Health Systems Performance Assessment, Background Document: Critical issues in health systems performance assessment. Washington, D.C.

WHO Regional Office for the Eastern Mediterranean (2001): Report on the Regional Consultation on the Conceptual Framework for Health System Performance Assessment. 9-7-2001, Ain Saadeh, Lebanon.

WHO Regional Office for Europe (2001): Report of the Regional Consultation on Health Systems Performance Assessment. 3-9-2001, Copenhagen.

WHO Regional Office for South-East Asia (2001): Report of the regional consultation and technical workshop on health systems performance assessment. 18-6-2001, New Delhi, India.

WHO Regional Office for the Western Pacific (2001): Report of the regional consultation on health system performance. 3-7-2001, Manila, Philippines.