



**MEASURING CLIENT-CENTERED
HEALTH CARE USING THE UNIVERSAL
WORLD HEALTH ORGANIZATION
CONCEPT OF “HEALTH SYSTEM
RESPONSIVENESS”**

Methods and applications

Nicole B. Valentine

Measuring client-centered health care using the universal World Health Organization concept of “health system responsiveness”

Methods and applications

Nicole B. Valentine

Measuring client-centered health care using the universal World Health Organization concept of “health system responsiveness” – methods and applications

PhD thesis, Erasmus University Rotterdam, The Netherlands

Cover design: L'IV Com Sàrl, Villars-sous-Yens, Switzerland

Layout: David de Groot, persoonlijkproefschrift.nl

Printing: Ridderprint BV, www.ridderprint.nl

Publisher: Ridderprint, Ridderkerk the Netherlands

ISBN/EAN: 978-94-6375-137-7

This thesis is partly realised due to the financial support of the Department of Public Health, Erasmus MC, University Medical Centre Rotterdam.

Copyright © 2018

Nicole Britt Valentine, Geneva, Switzerland, valentine_nicole@hotmail.com

All rights reserved. No part of this thesis may be reproduced in any form or by any means without written permission from the author.

**Measuring Client-Centered Health Care Using the Universal
World Health Organization Concept of “Health System
Responsiveness”**

Methods and applications

**Het meten van cliëntgerichte kwaliteit van zorg volgens het
universele WHO concept “responsiveness”**

Methoden en toepassingen

Thesis to obtain the degree of Doctor from the Erasmus University Rotterdam by
command of the rector magnificus

Prof. dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Friday December 14, 2018 at 9:30 hrs

by
Nicole Britt Valentine
born in Cape Town, South Africa



Doctoral committee

Promoter: Prof. dr. G.J. Bonsel

Other members: Prof. dr. E.W. Steyerberg
Prof. dr. E.K.A. van Doorslaer
Prof. dr. A. Franx

Dedication

To my father, Donald Geoffrey Valentine (RIP, 22 November 1931 - 22 January 2015), and my brother, Kim Geoffrey Valentine (RIP, 28 August 1957 – 1 August 1995 and 1 August 2016)

This is a brief life, but in its brevity it offers us some splendid moments, some meaningful adventures.

--Rudyard Kipling, Kim

CONTENTS

| | Page |
|---|------|
| List of abbreviations | |
| Chapter 1 Introduction | 11 |
| Part I Measuring responsiveness through household questionnaires | |
| Chapter 2 Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness | 25 |
| Chapter 3 Health systems responsiveness: a measure of the acceptability of health-care processes and systems from the user's perspective | 57 |
| Chapter 4 Health systems responsiveness and reporting behaviour: multilevel analysis of individual-level factors contributing to reporting behaviour heterogeneity in 64 countries | 97 |
| Part II Explaining why responsiveness matters for people, services and policy | |
| Chapter 5 Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of 'health systems responsiveness' in 41 countries | 125 |
| Chapter 6 What explains users' reports on health system responsiveness? Exploring health-care service and personal characteristics from surveys in 49 countries | 145 |
| Chapter 7 Exploring models for explaining the role of health systems responsiveness and social determinants in explaining universal health coverage and health outcomes: cross-sectional analyses of 57 countries | 171 |
| Part III Using responsiveness measures in the Netherlands' sub-system of perinatal care | |
| Chapter 8 Validity of a questionnaire measuring the World Health Organization concept of health system responsiveness with respect to perinatal services in the Dutch obstetric care system | 205 |

| | | |
|--|--|-----|
| Chapter 9 | Quality of perinatal care services from the user's perspective: a Dutch study applies the World Health Organization's responsiveness concept | 235 |
| Chapter 10 | General discussion | 259 |
| Part IV Additional bibliographic matter | | |
| | English Summary | 279 |
| | Dutch Summary | 291 |
| | Bibliography | 305 |
| | Authors and affiliations | 323 |
| | Manuscripts | 325 |
| | Manuscripts related to the thesis | |
| | Other related manuscripts | |
| | Other papers (unrelated, Pubmed) | |
| | Curriculum Vitae | 333 |
| | PhD portfolio | 337 |
| | WORD OF THANKS | 341 |
| | ANNEX | 343 |
| | Annex A. WHO responsiveness survey development | 345 |
| | A1. WHO responsiveness survey countries in this thesis | |
| | A2. WHO responsiveness household surveys characteristics | |
| | A3. WHO responsiveness domain experience questions | |
| | Annex B. WHO responsiveness questionnaires | 359 |
| | B1. Multi-Country Survey Study Questionnaire on Responsiveness | |
| | B2. World Health Survey Questionnaire on Responsiveness | |

LIST OF ABBREVIATIONS

| | |
|--------|--|
| AHRQ | United States Agency for Healthcare Research and Quality |
| CAHPS | Consumer Assessment of Health Plans Study / Providers and Systems |
| HDI | Human Development Index |
| MCS(S) | Multi-Country Survey Study on Health and Health Systems Responsiveness |
| OECD | Organisation for Economic Co-operation and Development |
| PREM | Patient reported experience measure |
| QUOTE | Quality Of care Through patients' Eyes |
| SD | Standard deviation |
| UHC | Universal health coverage |
| WHO | World Health Organization |
| WHS | World Health Survey |

CHAPTER 1

Introduction



AIM

The aim of this thesis is to provide a scientific evidence base on theoretical and empirical merits of the World Health Organization (WHO)'s "health system responsiveness" concept. The concept was part of the WHO's ambitious global measurement project on health systems' functioning. To quantify functioning of health systems globally, one needs universal and comparable metrics of health and other attainment variables, allowing within-country and cross-country comparisons.

The WHO developed a comprehensive measurement approach (including responsiveness) and launched its application with the production of the memorable 2000 World Health Report.^{1,2} The health system performance metrics presented in that report included the client-centredness of health services, termed "health system responsiveness".

The approach to measuring responsiveness followed that of a normal health survey, consisting of domains and items (questions) measuring performance levels on specified issues. Altogether 8 domains were covered, closely linking responsiveness to the United States Agency for Healthcare Research and Quality (AHRQ)³ Consumer Assessment of Health Plans Survey (CAHPS) questionnaire. The WHO implemented two rounds of multi-country household surveys which included the newly developed interviewer-supported responsiveness questionnaire: the Multi-Country Survey (MCS)⁴ and the World Health Survey (WHS)⁵, covering responsiveness measurement for 70 and 71 surveys across all modes, amounting to 106 interviewer administered surveys on responsiveness (see Annex A).

WHAT IS RESPONSIVENESS?

Eight domains are supposed to cover the most pertinent aspects of the client-health provider interactions, four "client orientation/setting" domains (choice, prompt attention, quality of basic amenities, social support (access to)); and four "respect for persons/personal" domains (autonomy, communication, confidentiality, dignity). In the WHO health system performance assessment framework, the responsiveness concept is one of three 'universal' health system measures or indicators: loss of health ('burden of disease') expressed in disability adjusted life expectancy, responsiveness, and fairness in financial contribution (Figure 1.1). Financing and responsiveness were weighted similarly in overall performance estimations, but health (loss) received a higher weighting.

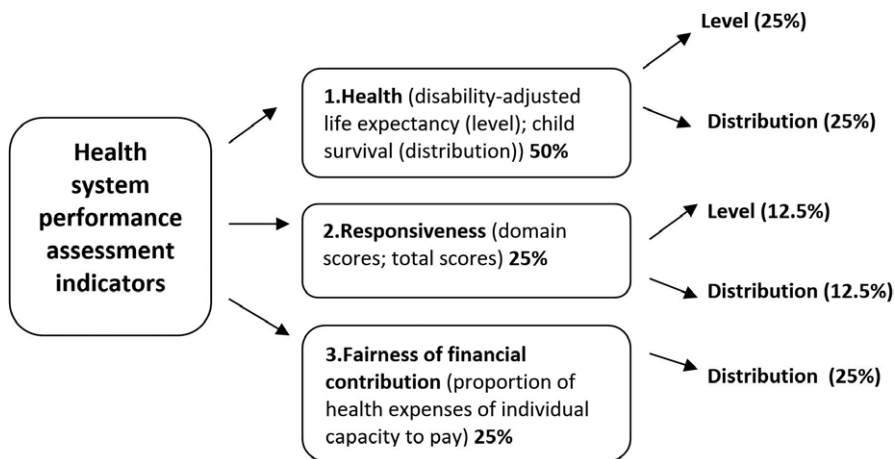


Figure 1.1 Health System Performance Assessment Framework

Source: Adapted by the author from WHO (2000)¹

SCOPE OF SERVICES FIT FOR RESPONSIVENESS MEASUREMENT

The scope of measurement of responsiveness can be broad: any organized health care or preventive action can be subject to assessment (discussed further in **chapter 3**), such as:

1. ambulatory care in response to acute needs;
2. ambulatory care for chronic conditions;
3. inpatient care for short-term stays (typically >24 hours, <3 months);
4. long-term institutionalized care: e.g., for populations with mental illnesses, disabilities related to physical health conditions or elderly populations;
5. non-excludable public health interventions: e.g., public health promotion for communities or population groups such as access to improved water and sanitation, smoking bans;
6. opportunities for participation in health system governance: e.g., shaping the health system and issues affecting health;
7. administrative and financial transactions: e.g., ease of making payments, obtaining prescriptions for chronic medication, receiving reimbursement from insurance.

The unit of aggregation (e.g., community, hospital, national, scheme) too is not fixed. Rather it is guided in analysis by the envisaged unit of accountability for responsiveness (e.g., local government, a particular provider, national department of health, or insurance company).

HISTORICAL RECEPTION OF THE THREE WHO PERFORMANCE INDICATORS, AND OF RESPONSIVENESS IN PARTICULAR

Since 2000, the concepts and derived indicators of the WHO 2000 World Health Report have shared different fates with respect to their absorption into inter-governmental accountability frameworks and the academic world of health services research and public health. By 2015, the *burden of disease* concept and the related use of DALY's as outcome measure, was well-absorbed in science and, to a large extent, into the larger health policy arena, playing *inter alia* critical roles worldwide in prioritization of health care packages of drugs. It was readily accepted for monitoring specific diseases, but not for monitoring overall life expectancy, as testified by the indicator framework⁶ related to the United Nations General Assembly resolution on the Sustainable Development Goals (SDGs).⁷

Financial protection coverage, derived from the original WHO 2000 Health Systems Performance Assessment framework, emerged as an important axis of performance assessment in the monitoring of universal health coverage (UHC) promoted by the World Health Organization and the World Bank.^{8,9} UHC financial protection is optimal if all people who need services, use them without financial hardship. The failure of coverage is measured using rates of catastrophic expenditure.

By comparison, it is fair to say that the third key concept, responsiveness, received less immediate and consistent acclaim. Yet almost 20 years later, Patient Reported Experience Measures (PREMs) have developed as a latter heritage. At international levels, responsiveness appears in the indicator framework of WHO¹⁰ and the OECD¹¹ where data on a few countries are available in the Health Care Quality Indicators repository. At the time of writing (May 2018), the OECD has data on patient experience for prompt attention, communication and autonomy domains (see: <http://www.oecd.org/els/health-systems/hcqi-responsiveness-and-patient-experiences.htm>). Following a hiatus coinciding with the backlash to the World Health Report¹², national-level implementation of responsiveness measurement has been pursued in some countries, notably in the United Kingdom¹³, in the United States¹⁴, but also in Australia¹⁵ and in the Netherlands¹⁶. In the Netherlands, the national insurance stakeholders rely on the Consumer Quality Index (CQ-Index) measures for performance measurement of sub-systems of care which are intimately related to the responsiveness concept.¹⁷

WHY ABSORPTION OF RESPONSIVENESS WAS DELAYED

Novelty is not a factor distinguishing responsiveness from the burden of disease and financial protection concepts, as all were new concepts. But perhaps the nature of the change with past concepts for responsiveness was more pronounced in several ways.

The WHO responsiveness concept was explicitly derived from the Donabedian work on quality of care^{2,18}, but there are many differences. Responsiveness has a central focus on non-clinical aspects of the care process, thus separating quality in result (health outcome, the decrease in burden of the disease) from quality in the service's client orientation. This separation departs from the more familiar landscape of quality. Responsiveness thus derived an understanding of what is 'good' service from non-clinical theoretical underpinnings in human rights in particular, but also in consumer theory, medical ethics and legal instruments governing communication and decision making. The replacement of Donabedian's 3-tier building blocks (structure, process, outcome) with measurement on different levels (e.g., organisation, presence of legal rights, performance of procedures) by a single measurement principle (ask the client/user/patient about his/her experiences), was an empirical transformation. Service quality and client orientation following the responsiveness philosophy cannot be more than or beyond what people actually perceive.

These changes implied the formidable task of creating a 'universal' measurement tool ('quality experience surveys'), but also had major advantages: 1) the processing and analysis of individual experiences can follow the clinical outcome framework, including inequality measures; 2) paired analytical designs can determine how lack of service quality affects health outcomes, and the reverse, how severity of disease may limit attainable levels of service quality; and 3) discussions of a good versus bad system based on ideology were removed from this empirical table. A hospital was not 'good' based on its number of service desks, but by the experienced waiting times by users. A care provider was not 'good' in terms of client oriented quality by his/her professional degree, but by experienced dignity and effective communication. Actually responsiveness, as with patient satisfaction, was an early expression of the now more familiar 'patient' and 'people'-centred care 'movement'.

The development of the measurement tool was based, as mentioned previously, on theory-consistent 'universal' questions addressed to individuals, covering the whole spectrum of personal and setting quality aspects, and which allowed for case-mix adjustment and the handling of expectation bias (if present). Not only were these changes a radical scientific and policy change, but they also affected a whole industry of consultancy and IT stakeholders increasingly dedicated to measuring patient experience over the past decade.

The apparent initial reluctance to accept a universal client-based quality concept may in part have arisen from the inevitable cultural and political nature of health services, as compared with health outcomes or, to a lesser extent, financing. The claim that services worldwide, despite their diversity, could simply be compared by one unified neutral set of measurable criteria, ignored the country-related ideological conceptions of 'good care'. Such non-ideological approaches have winners and losers. There was initially a similar reluctance to accept the burden of disease concept which also created new winners and losers. But in that case there were already similar well-accepted concepts, notably the quality-of-life years (QALY) and the (disease-free) life expectancy concept.

DELAYED BUT NOT HALTED

The more one considers the global context of international human rights, globalization and international migration, the more that the rising expectation is understandable: that health services¹⁹, like any other paid service, should be made accountable with a common assessment framework. It was for long clear that against the political background, the convergence of measures and systematic use of responsiveness-like data would not happen spontaneously. The early WHO work set the standard for later initiatives.

It took time. But with the current routine introduction of patient-reported outcome measures (PROMs) and PREMs in quality frameworks and clinical registries, as well as international initiatives like the International Consortium for Health Outcomes Measurement, it is reasonable to conclude that the 'patient' movement has made the case for WHO's responsiveness. Added to this, the WHO has increasingly (2013-15) focused international attention on redefining 'universal health coverage' (UHC) as including coverage with needed quality services (without financial hardship). As the set of UHC services are defined, so policy-makers and populations will be increasingly sensitive to quality (clinical and non-clinical) dimensions of service delivery.

THE THESIS DATA

WHO surveys

This thesis focuses on measuring responsiveness for clients with experiences in either outpatient or inpatient services (public or private).

The responsiveness questionnaire modules WHO fielded in the 2000-01 MCS and in the 2002-04 WHS covered 70 and 71 surveys of which, 41 interviewer and 65 interviewer administered surveys, respectively, representing 83 countries (excluding overlaps, see Annex A), were analysed for this thesis. The full questionnaires comprised modules on socio-de-

mographic background, social capital, own health, own health care utilization, own responsiveness experience, and health and responsiveness vignettes (for the full questionnaires: <http://www.who.int/healthinfo/survey/en/>; the responsiveness module / questionnaire: Annex B). Approximately 105,806 respondents in the MCS and 152,445 respondents in the WHS, totalling approximately 258,000 records, answered questions related to responsiveness in 106 interviewer administered surveys.

Questions focused on the performance of the service in the client's experience (e.g., "how would rate your experience of the way health care providers communicated with you?"); the importance of domains, or "preference" of clients (e.g., "how important is "clarity of communication" to you? This means having the provider listen to you carefully; having the provider explain things so you can understand; having time to ask questions"); or 'vignettes' questions – hypothetical scenarios describing the quality of interactions with health service providers, which respondents were asked to evaluate. The WHO long questionnaires consisted of 28 questions on domain experiences, with ordinal verbal response categories: 13 for outpatient and 15 for inpatient questionnaires; and 8 importance-of-domain questions. The short questionnaire consisted of 15 domain experience questions; 7 outpatient; 8 inpatient (with the same importance questions and a reduced set of vignette questions).

The ReproQ survey

The ReproQ was developed between October 2009 and February 2010 by adapting the WHO responsiveness questionnaire items to the perinatal care context. Records for 171 women women who participated in the survey were analysed for this thesis.

The ReproQ questionnaire was developed to assess the responsiveness outcomes of perinatal health care system in the Netherlands and is based on the same 8 domains identified in WHO's review, i.e. dignity, autonomy, confidentiality, communication, prompt attention, social consideration (labelled initially as *Access to Social Support* or *Access to Family and Community Support*), quality of basic amenities, and choice ("and continuity"). The ReproQ asked the same questions for the three phases of perinatal care: antenatal phase (the period from the onset of pregnancy until the onset of delivery), birth phase (actual delivery) and post partum phase (covering the first 10 days after childbirth). Constructing parallel questionnaires for antenatal and postnatal care separately, the ReproQ consisted of 104 questions on responsiveness experiences (25 antenatal, 40 birth, 39 postpartum phase), 29 questions for maternal and health care characteristics and 8 importance-of-domain questions.

Ethical study approval is reported for the MCS as obtained from the WHO Sub-Committee for Research Involving Human Subjects; for the WHS from the Harvard School of Public Health's

Institutional Review Board as well as from the relevant ethics committee in different survey sites; and for the ReproQ, from the Medical Ethical Committee, Erasmus Medical Centre, Rotterdam, the Netherlands. In all cases, respondent consent was obtained before interviewing.

THE THESIS STUDY QUESTIONS

The thesis is divided into three sections according to the main themes and leading study questions it addresses.

The first part of the thesis addresses psychometric testing of the WHO household survey data on responsiveness and innovative analyses on the nature and causes of individual-level reporting behaviour biases. It complements the published approaches using the HOPIT²⁰ model. The second part applies the responsiveness measures to global and within-country health system comparisons and focusses on exploring linkages to health system policies.

The third part of the thesis addresses the application of responsiveness measurement and performance reporting to the sub-system of perinatal care in a single country's setting in the Netherlands.

The seven leading study questions are grouped below under each of the main thematic parts of the thesis.

PART I: Measuring responsiveness through household questionnaires

1. Do populations across different countries and from different socioeconomic strata within countries share a common understanding of health system responsiveness domains? (chapters 2, 3)
2. Which characteristics of individuals affect reporting behaviour biases when using the responsiveness domain question and answer format, and how? (chapter 4)

PART II: Explaining why responsiveness matters for people, services and policy

3. Which domains of responsiveness are more valued and by whom? (chapter 5)
4. Which health service characteristics drive responsiveness performance levels and are performance measures equity-sensitive? (chapter 6)
5. How is responsiveness considered to influence other important health system outcomes, like service coverage ('access') and health ('clinical' or health outcomes)? (chapters 3, 7)

PART III: Using responsiveness measures in the Netherlands' sub-system of perinatal care

6. Can responsiveness measures used in general household questionnaires be applied to measure the quality of a specific sub-system of care? (chapter 8)
7. Which personal, health-related experiences are most associated with responsive performance? (chapter 9)

The thesis aims to offer the reader a testable and critical account of the performance of the proposed WHO concept with regard to its measurement.

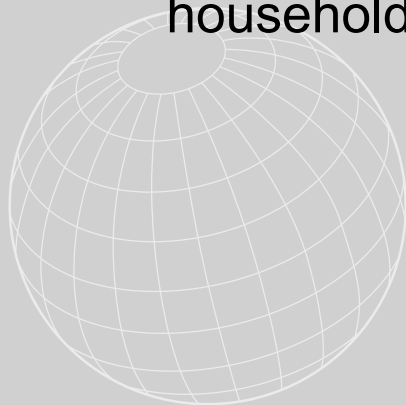
REFERENCES

1. World Health Organization. Health systems: improving performance. The world health report 2000. Geneva: World Health Organization; 2000.
2. Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bull World Health Organ.* 2000;78(6):717–31.
3. Hargraves JL, Hays RD, Cleary PD. Psychometric properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 adult core survey. *Health Serv Res.* 2003;38(6 Pt 1):1509–27.
4. Üstün TB, Chatterji S, Villanueva M, Bendib L, Celik C, Sadana R, Valentine NB, Ortiz J, Tandon A, Salomon J, Cao Y, Wan Jun X, Özaltın E, Mathers C, Murray CJL. WHO Multi-Country Survey Study on Health and Responsiveness 2000-2001. Global programme on evidence for health policy. Discussion paper no. 37. Geneva: World Health Organization; 2001.
5. Üstün TB, Chatterji S, Mechbal A, Murray CJL, WHS Collaborating Groups. The World Health Surveys. Chapter 58. In: Murray CJL, Evans DB, editors. *Health Systems Performance Assessment: debates, methods and empiricism.* Geneva: World Health Organization; 2003. p.797–808.
6. United Nations Economic and Social Council: Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators. In: E/CB3/2016/2/Rev1. Edited by Commission S. New York: United Nations; 19 February 2016.
7. United Nations General Assembly: Transforming our world: the 2030 Agenda for Sustainable Development (Resolution 70/1). New York, NY: United Nations; 2015.
8. World Health Organization, World Bank. Tracking universal health coverage: first global monitoring report. Geneva: WHO, 2015.
9. World Health Organization, World Bank. Monitoring progress towards universal health coverage: framework, measures and targets. Geneva: World Health Organization; 2014.
10. World Health Organization, International Health Partnership Initiative. Monitoring, evaluation and review of national health strategies: A country-led platform for information and accountability. Geneva: WHO; 2011.
11. Kelley E, Hurst J. Health care quality indicators project conceptual framework paper. Organisation for Economic Co-operation and Development (OECD) Working Papers 2006. Paris: OECD; 2006.
12. Almeida C, Braveman P, Gold M, Szwarcwald CL, Ribeiro JM, Americo M. Methodological concerns and recommendations on policy consequences of the world health report 2000. *Lancet.* 2001, 257:1692–7.
13. Department of Health: NHS. Outcomes Framework for England, 2013-2016. London: Department of Health, United Kingdom; (January, 23) 2013.
14. Agency for Healthcare Research and Quality. 2014 National healthcare quality and disparities report. Rockville MD: Department of health and human services, Agency for Healthcare Research and Quality; (May) 2015.

15. Council of Australian Governments (GOAG). National Health Reform: performance and accountability framework. Canberra; 2015. Available from: <https://www.aihw.gov.au/getmedia/ea9b2361-38de-43f3-9426-8705fcc8f1da/performance-and-accountability-framework.pdf.aspx>
16. National Institute for Public Health and Environment. Dutch health care performance report 2014 (in Dutch: Zorgbalans 2014. De prestaties van de Nederlandse gezondheidszorg). Editors, van den Berg MJ, De Boer D, Gijzen R, Heijink R, Limburg L, Zwakhals S. Bilthoven: National Institute for Public Health and Environment (RIVM); 2015.
17. Delnoij DMJ, Rademakers JJ, Groenewegen PP. The Dutch Consumer Quality Index: an example of stakeholder involvement in indicator development. *BMC Health Serv Res.* 2010; 10:88–88.
18. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
19. Blendon RJ, Schoen C, DesRoches CM, Osborn R, Zaper K. Common concerns amid diverse systems: health care experiences in five countries. *Health Affair.* 2003(May-Jun);22(3):106–21.
20. Rice N, Robone S, Smith P. Vignettes and health systems responsiveness in cross-country comparative analyses. *J R Stat Soc Ser A.* 2012;175(2):337–69.

PART I

Measuring responsiveness through
household questionnaires



CHAPTER 2

Measuring quality of health care from the user's
perspective in 41 countries: psychometric
properties of WHO's questions on health systems
responsiveness



N.B. Valentine, G.J. Bonsel, C.J.L. Murray
Quality of Life Research. 2007; 16(7):1107-25

ABSTRACT

Objective. To evaluate, for different populations, psychometric properties of questions on “health systems responsiveness”, a concept developed by the World Health Organization (WHO) to describe non-clinical and non-financial aspects of quality of health care.

Data sources/study setting/data collection. The 2000–2002 WHO Multi-Country Study comprised 70 general population surveys. Forty-one surveys were interviewer-administered, from which we extracted respondent records indicating ambulatory and inpatient health services use (excluding long-term institutions) in the previous 12 months (50,876 ambulatory and 7,964 hospital interviews).

Study design. We evaluated feasibility, reliability, and construct validity of using 33 items with polytomous response options, comparing responses from populations identified by countries, sex, age, education, health and income.

Principal findings. Average item missing rates ranged from 0 to 16%. Domain-specific alpha coefficients exceeded 0.7 in 7 (of 9) cases. Average intertemporal reliability was acceptable in 6 (of 10) sites, where Kappas ranged from 0.54 to 0.79, but low in 4 sites ($K < 0.5$). Kappa statistics were higher for male, educated and healthier populations than for female, less educated and less healthy populations. Factor solutions confirmed the domain structure of 7 domains (only 7 were operationalized for ambulatory settings). As in other studies, higher incomes and age was associated with more positive responsiveness reports and ratings.

Conclusion. Quality issues addressed by WHO’s questions are understood and reported adequately across diverse populations. More research is needed to interpret user-assessed quality of care comparisons across population groups within and between countries.

INTRODUCTION

The quality of users' interactions with health services are intrinsically and instrumentally important to quality of life outcomes. Yet few international agencies have undertaken extensive studies of quality of health care from the user's perspective.¹ This made the World Health Organization's (WHO) proposal in 2000 to develop a universal, population-level indicator called "health systems responsiveness" a pioneering step. The proposed concept covered a set of non-clinical and non-financial dimensions of quality of care that reflected respect for human dignity and interpersonal aspects of the care process, which, as Donabedian remarked, "is the vehicle by which technical care is implemented and on which its success depends".²

WHO formed a technical collaboration agreement with the United States Agency for Healthcare Research and Quality (AHRQ) to develop a questionnaire to measure health systems responsiveness. Under the auspices of the Multi-Country Survey Study on Health and Health Systems Responsiveness (the MCS Study), the questionnaire was administered between 2000 and 2002. This paper presents the first evaluation of the feasibility, reliability and validity of responsiveness questions used in the MCS Study. In the best-case scenario, responsiveness questions would have good psychometric properties and differ little by the characteristics of individual respondents or country of administration.

METHODS

Literature review and defining the responsiveness domains

The responsiveness concept was based on literature in the fields of medical ethics, human rights, and human development, and identified aspects of health care delivery important to users apart from health outcomes.³ Electronic literature searches conducted by a WHO consultant between July and November 1999, using Medline, Psychlit, and Social Science Citation Index databases, covered literature published between 1990 and 1999. Important search terms were "quality of care", "dignity", "confidentiality", and "choice". The search term, 'patient satisfaction', while implying a different measure, was also used because it covered important domains of users' experiences. Retrieved literature included seminal articles such as Thompson and Sunol⁴, Sitzia and Wood⁵, and Wensing et al.⁶ Articles frequently cited in bibliographies (more than 3 times) but published before 1990 were also extracted (e.g., Ware & Hays⁷), as were relevant questionnaires like the Consumer Assessment of Health Plans Study (CAHPS®) Questionnaire (now the Consumer Assessment of Healthcare Providers and Systems), the Picker Patient Experience Questionnaire, the Patient Satisfaction Questionnaire, and the QUality Of care Through patients' Eyes (QUOTE) Questionnaire.

Themes covered in these questionnaires echoed Donabedian's² concept of interpersonal quality of care as well as other aspects important for the acceptability of care.⁸ Themes were divided into 8 internally homogenous and comprehensive domains describing outcomes of the care process apart from positive health outcomes and non-improvement: dignity, autonomy, confidentiality, communication, prompt attention, quality (of) basic amenities, users having access to social support networks during treatment (labelled 'social support'), and choice (of health care providers).⁹ Operationalizing the concept followed Parasuraman et al., who identified respondents' judgments of service quality as different from 'satisfaction' measures.¹⁰ 'Satisfaction' was seen as more closely associated with hearsay, impressions, and comparisons of expectations with actual experiences while experience judgments were more closely associated with objective service realities.¹¹

Testing responsiveness questions

Three field tests shaped the final MCS responsiveness questionnaire developed by the WHO team, whose membership included two of the authors to this paper. In 1999, the first survey sampled 'key informants', that is, professionals or researchers, rather than the general population across 35 countries (n=1791). Survey investigators, chosen for their expertise to lead the surveys in each country and assembled by WHO to discuss the results, supported the AR-HQ-proposed inclusion of communication as a distinct domain (instead of subsumed under dignity and autonomy). The general population surveys also in 1999 (n=450 across 3 countries) and in 2000 (n=811 across 8 countries) showed that psychometric properties of the responsiveness questions were adequate (e.g., missing <3%, Kappa (K) ≥ 0.6). Cognitive interviews accompanying the 2000 survey (n=174) suggested that key concepts (e.g., dignity) held equivalent meanings in diverse languages, including Chinese, Egyptian Arabic, and Slovakian.

The MCS Study questionnaire and responsiveness 'module'

The MCS Study questionnaire came in a 'short' and 'long' form, of which the responsiveness 'module' was one component. Other modules covered health and socio-demographics. The long questionnaire, containing 9 modules, was used in only 12 countries.ⁱ

The responsiveness module in the long questionnaire contained 127 responsiveness items (20 to 25 minutes to administer) and 87 items in the short version of the questionnaire (15 to 20 minutes to administer). The difference in responsiveness items was mostly due to extra sections on home care (23 items) and utilization (13 items) (e.g., receiving medication). The responsiveness module had three components: polytomous-scaled 'performance' questions (judgments of experiences); importance questions (ranks of domain importance); and 'expectations' questions (expectations regarding treatment standards). Appendix 2.1 contains the full wording of the

performance questions, and, www.who.int/responsiveness/surveys/en, the full questionnaire. This paper focuses on the performance questions.

Responsiveness performance questions

Responsiveness performance questions covered ambulatory (22 items) and inpatient (11 items) visits (defined as an overnight stay of 24 hours or more). Eight items came from the CAHPS-2.0 Adult questionnaire.¹² If a respondent used both ambulatory and inpatient services in the previous 12 months, they answered questions on the same domains for both these encounters (except in social support, which was only in the inpatient section). Item handles are listed in Table 2.1. All questions used similarly ordered 4-point (always, usually, sometimes, never) or 5-point (mainly: very good, good, moderate, bad, very bad) verbal response options, alternatively known also as 'report' or 'rating' scales. To reduce the length of the questionnaire, a decision was taken to have a shorter inpatient section, by reducing the number of items per domains.

MCS Study countries and survey administration

The MCS Study questionnaire was administered by governmental agencies, universities and survey companies. Study protocols and processes were cleared by the WHO Sub-Committee for Research Involving Human Subjects and respondent consent was sought before interviewing.¹³ One-hundred-and-forty-one thousand interviews were completed through 41 interviewer-administered surveys and 29 self-administered surveys. This represented a study participation rate of 75%, calculated by dividing the total number of attempted contacts by the number of effective contacts (see Appendix 2.1 for response rates). To remove possible confounding associated with administration mode, and for reasons of space, this paper focuses on the 41 interviewer-administered surveys in 41 countries (see Appendix 2.2) (also with a participation rate of 75%).

A detailed translation protocol required forward and back translation of key terms by a third person, and an expert panel review (see underlined phrases in Table 2.1). One to 3 national languages were used per country. Translated questionnaires were tested on 20 to 100 local respondents. Sampling schemes used stratified multi-stage random sampling or cluster sampling with random walk, and sampling frames such as recent censuses. Surveyors aimed for national representation, except in India, China, and Nigeria, where surveyors aimed for samples to represent the populations of the few conveniently selected provinces (or states). Interviewers called on households between 2 and 10 times. Within households, eligible respondents (18 years or older) were selected using the 'most recent birthday' method or Kish tables. Further details of the Study's administration are described elsewhere.¹³

Data cleaning and selection

Cleaning procedures applied to the MCS Study dataset checked that numbers assigned to verbal response options were consistent in translated questionnaires. Missing data were also completed if information was available elsewhere (e.g., the household roster). From the cleaned dataset, we extracted all records reporting health service use in either or both ambulatory or inpatient setting in the previous 12 month and kept those where the summary question in 4 or more responsiveness domains and the self-report question asking about service utilization were completed (99.9%). This process yielded a dataset containing 105,806 respondents, of whom 56% were classified as 'users'. Analyses were performed with Stata Special Edition v7. Inappropriate missing rates for bivariate analysis consisted of the combined missing rates of both variables. For multivariate analyses, we completed missing data using the maximum likelihood method specified in NORM v2.03 ('Norm' procedure) with multiple imputations.^{14,15}

Variable coding

We coded verbal response options for the responsiveness questions to numeric values, with 1 corresponding to the worst, and 4 or 5 to the best response options. Answers of "refuse", "don't know", or "not applicable" were recoded to missing (<1%). While the items were strictly ordinal-level, we treated them as interval-level. Report and rating values were treated as quasi-cardinal, as is common in user-evaluated research.¹⁶

Other variables were coded as follows. We took the country variable as categorical and to represent culture. To describe development context, we used the Human Development Index (HDI) as categorical, condensing 3 HDI categories into 2 (more (high HDI) and less (low HDI) developed - see Appendix 2.2) (United Nations Development Programme's (UNDP's)¹⁷). Population subgroups within countries were distinguished in terms of sex (male or female), age (<56 years, >55), education (<8 years, >7 years), and self-assessed health (a 5-point scale classified as 'healthy' ("very good" and "good") and 'less healthy' ("moderate", "bad", "very bad")). Additional analyses were run using more refined age groups (5 year intervals from 18 to 85) and education categories (from 0 to 20, and >20). For one validity analysis, age was split into 3 categories (≤35 yrs, 36-55 yrs, >55 yrs). An income quintile variable from the survey was used in one of the construct validity analyses.

Psychometric tests

We used a standard set of feasibility, reliability (internal consistency and temporal) and validity tests. We investigated the responsiveness questions' psychometric properties for the sample as a whole (the pooled dataset), for groups of countries classified as more and less developed according to the HDI, and for differences subgroups. A parsimonious set of results are reported. Additional results are available in appendices.

Feasibility tests used survey response rates, respondent inappropriate missing rates, item missing rates (3% cut-off), response frequencies, ceiling effects, and item mean rankings. Ceilings effects higher than 50% of respondents with the most positive response, were considered unproblematic if not present in all questions in a domain or across all countries. According to the literature¹⁸, similar rankings of item means within domains for similar populations indicate that the translation process has left unchanged the relative ordering of items within domain scales. We compared the relative ranking between paired combinations of item means within each domain for each country to a 'standard' set by the ranking for the majority (50% or more) of the countries.

Scale internal consistency, a measure of *reliability*, was assessed with inter-item and item-rest standardized correlation coefficients (Pearson correlation coefficient (r) >0.40 ^{18,19}). We expected higher inter-item and item-rest correlations between items in the same domain. Amidst varying standards in the literature, we chose >0.80 for good alpha coefficients and <0.70 for suboptimal (Nunnally and Bernstein²⁰ indicate that 0.7 is "acceptable").

Item temporal reliability was assessed with weighted Kappa statistics, which were judged as modestly reliable if between 0.41 and 0.60.²⁰ Test-retest results were available for 10 countries, which all used the same questionnaire, from which there were 2,854 ambulatory and 417 inpatient retest interviews. Interviews were re-administered by the same interviewers between 8 days and 1 month after the initial interview. Two-by-two tables, using a Kappa cut-off value of 0.65, compared differences in Kappa statistics for paired population groups (e.g., older and younger) with Chi-square two-tailed of statistical significance ($p=0.01$).

Associations between the psychometric properties described by the statistics mentioned above were also assessed using correlation coefficients for the more refined age and education groupings. Associations were judged as moderate if correlations lay between 0.30 and 0.80.¹⁹

Assessing *content validity* involved the tasks described earlier, which included a literature review, discussions with principal investigators on the key informant surveys, and field tests, including cognitive interviews.

Assessing *construct validity* involved assessing the domain structure underlying the data using maximum-likelihood (ML) factor analysis for the ambulatory items only (most inpatient domains had only 1 item). We used Kaiser's eigenvalue rule (factors with eigenvalues greater than 1) to identify important factors and Cattell's scree test to visualize the eigenvalues. Kaiser's eigenvalue rule also stipulates that item loadings on factors need to be 0.40 or greater.^{20,21} The set of

items in the factor analyses excluded 2 skip pattern autonomy items in order to maintain the full 50,876 observations (versus 36,423).

Construct validity was also assessed using three hypotheses. One: higher responsiveness was associated with higher human development. Two: cross-country differences were assessed by comparing mean scores for more and less developed countries (t-tests) and correlating (Pearson) country responsiveness scores and HDI ranks (lower rank meaning less development). There would be higher responsiveness in wealthier populations. Three: older populations would report higher responsiveness.²²

Hypothesis two compared responsiveness and income for ambulatory services only (due to high missing rates for the income variable and low visits to inpatient services). The income quintile variable was only used in 29 countries where income missing rates were less than 15% (average missing, 9%) and completed with the 'Norm' procedure. Income was preferred over education in spite of its higher missing rate as it was more likely that richer people would have access to more responsive health care services than people with high education.

Composite responsiveness scores were calculated by averaging individual-level 0 to 1 scores within domains, then across domains up to the country level. Composite scores were recoded back to 1-5 categorical values for hypothesis three only (≤ 0.2 to 1, > 0.2 and < 0.4 to 2, etc. to 5) and associations were tested using Gamma (range: -1 to +1), a correlation coefficient for ordinal variables.

RESULTS

Respondent characteristics

From 105,806 respondents in 41 countries, there were 50,876 ambulatory and 7,964 inpatient interviews. In more developed countries, 52% of users were female; in less developed countries, 59%. Age (45 vs. 40 years) and education levels (10 vs. 7.5 years) reflected the demographic differences in development settings. About 55% of users in both development settings said their health was good or very good (which was below the average of 73% for non-users). Appendix 2.2 contains descriptive statistics for the Study samples.

Feasibility analyses

Response rates for the interviewer-administered surveys were on average 70% for effective contacts (11%-99%, $n=37$) and 46% for attempted contacts (10-84%, $n=29$). Ex-post comparisons of the survey sample's age and sex profiles with UN population statistics showed that in both sexes, younger respondents (<35 years) were under-represented and older respondents

(60–65 years) were over-represented (13). UN education statistics (averaging 8 years for the 41 countries) showed that most samples were biased towards more educated respondents (United Nations Educational, Scientific and Cultural Organization).²³

The average item missing rates (2%) are shown in Tables 2.1 and 2.2 (see Appendix 2.3 for inpatient items). Ten percent of respondents to ambulatory items had inappropriate missing responses—7% were missing 1 item (mostly, “using a provider other than your usual one” (55%)). Only 1% of respondents to the inpatient section had inappropriate missing items (80% of 1% was attributed to the ‘religion’ item, “practicing religious/traditional observances in hospital”). Countries shared similar item missing patterns, but, in 17 of 82 cases (2 (ambulatory or inpatient)*41), countries had higher absolute missing rates (>5%). It was interesting to note that for the ‘religion’ item, former Soviet countries (n=9) had higher average inappropriate missing rates (22%) than Islamic countries (n=10) (1%).

Item missing patterns were similar across populations defined by sex, age, education, and health, and rarely exceeded 3%, except in older respondents (3.4%) where modest positive correlations were observed for age groups defined by 5 year intervals ($r=0.50$ for ambulatory items and $r=0.50$ for inpatient items) and for education groups ($r=0.50$ for ambulatory items; $r=0.40$ for inpatient items). Correlations were positive except for education groups and autonomy item missing rates, which also had the highest correlation coefficients ($r=-0.80$).

There were no responses in the most negative categories of items in 32 of 902 (41*22) cases for ambulatory items, and in 114 of 451 cases for inpatient items. Fourteen (of 33) items had ceiling effects, but only 8 exceeded 60%. No domain displayed ceiling effects for all items or countries.

Translation equivalence was comparable for high human development countries to the 11 high development countries from the International Quality of Life Assessment (IQOLA) project.¹⁸ IQOLA’s questionnaire contained 6 health domains with subscales containing more than 1 item. For IQOLA 12% of countries had item mean rankings differing from the standard (taken as the item ranking for the majority ($\geq 50\%$) of countries) compared with 16% in the MCS Study. For the 17 low human development countries, this Figure was 22%.

Internal consistency reliability

Inter-item and item-rest correlations exceeded 0.4 (see Table 2.2). Correlations were higher between items within a domain. The highest inter-domain correlations were for ‘overall’ dignity and communication items (see Table 2.2, A6 and A10, $r=0.60$), and for ‘overall’ communication and autonomy items ($r=0.60$). Only alpha coefficients for social support (0.62) and prompt attention (0.65), were less than 0.75. The similarity of alpha coefficient patterns across 41 countries indicated that items corresponded with similar domains in different contexts (see Appendix 2.4).

Table 2.1 Likert-scaled responsiveness ambulatory items and item properties

| Item ^a | Item handles (see full question in Appendix 2.1 and full questionnaire at www.who.responsiveness/surveys/en) ^b | Response Options | Mean (0-1) | SD | Ceiling effect (%) | Missing rate (%) | Kappa (0-1) | Domain |
|-------------------|---|------------------|----------------------|------------------------------|--------------------|------------------|-------------|---------------------------|
| | | | Study dataset pooled | 41 countries averaged (ave.) | 10 countries (ave) | | | |
| A1 | getting care <u>as soon as you wanted?</u> | c | 0.81 | 0.21 | 55 | 1.2 | 0.62 | Prompt attention |
| A2 | <u>getting prompt attention at the health services</u> in the last 12 months? | d | 0.78 | 0.17 | 23 | 1.0 | 0.58 | |
| A3 | doctors (nurses or other health care providers) <u>treat you with respect?</u> | c | 0.88 | 0.17 | 64 | 0.3 | 0.63 | Dignity |
| A4 | office staff treat you with respect? | c | 0.86 | 0.19 | 59 | 2.3 | 0.60 | |
| A5 | <u>physical examinations and treatments done in a way that respected your privacy?</u> | c | 0.89 | 0.18 | 69 | 1.4 | 0.59 | |
| A6 | <u>getting treated with dignity</u> | d | 0.83 | 0.15 | 38 | 0.5 | 0.60 | Autonomy |
| A7 | doctors (nurses...) <u>listen carefully to you</u> | c | 0.86 | 0.18 | 58 | 0.4 | 0.55 | |
| A8 | doctors (nurses...) there, <u>explain things in a way you could understand</u> | c | 0.84 | 0.20 | 56 | 0.5 | 0.58 | |
| A9 | doctors (nurses...) give you time to <u>ask questions about your health problem or treatment</u> | c | 0.82 | 0.22 | 53 | 0.8 | 0.54 | |
| A10 | <u>how well health care providers communicated with you</u> | d | 0.76 | 0.26 | 33 | 0.5 | 0.59 | Communication |
| A11 | <u>involve you as much as you wanted to be in deciding about the care</u> | c | 0.75 | 0.28 | 45 | 2.4 | 0.66 | |
| A12 | <u>ask your permission</u> before starting tests or treatment | c | 0.83 | 0.23 | 46 | 2.7 | 0.66 | |
| A13 | your experience of getting involved in making decisions about your care or treatment as much as you wanted | d | 0.87 | 0.19 | 27 | 3.0 | 0.63 | Confidentiality |
| A14 | talks with your doctor done privately so other people <u>could not overhear what was said?</u> | c | 0.81 | 0.16 | 61 | 2.2 | 0.59 | |
| A15 | doctor (nurses...) keep your personal information confidential | c | 0.77 | 0.18 | 69 | 10.0 | 0.57 | |
| A16 | <u>health services kept information about you confidential</u> | d | 0.83 | 0.16 | 43 | 7.0 | 0.61 | |
| A17 | <u>to get to a health care provider you were happy with?</u> | e | 0.89 | 0.19 | 69 | 2.9 | 0.65 | Choice of (care) provider |
| A18 | <u>using other health care services other than the one you usually went to?</u> | e | 0.89 | 0.19 | 69 | 16.2 | 0.65 | |
| A19 | <u>being able to use a health care provider or service of your choice</u> over the last 12 months? | d | 0.80 | 0.18 | 34 | 3.4 | 0.60 | |

Table 2.1 Likert-scaled responsiveness ambulatory items and item properties (continued)

| Item ^a | Item handles (see full question in Appendix 2.1 and full questionnaire at www.who.responsiveness/surveys/en) ^b | Response Options | Mean (0-1) | SD | Ceiling effect (%) | Missing rate (%) | Kappa (0-1) | Domain |
|-------------------|---|------------------|----------------------|------------------------------|--------------------|------------------|-------------|--|
| | | | Study dataset pooled | 41 countries averaged (ave.) | 10 countries (ave) | | | |
| A20 | the basic quality of the waiting room, for example, space, seating and fresh air. | d | 0.77 | 0.17 | 25 | 0.7 | 0.65 | Quality of basic amenities |
| A21 | the cleanliness of the place? | d | 0.79 | 0.16 | 31 | 0.8 | 0.64 | |
| A22 | the overall <u>quality of the surroundings</u> , for example, space, seating, fresh air and cleanliness | d | 0.77 | 0.16 | 25 | 0.8 | 0.65 | |
| I1 | getting attention from doctors as quickly as you wanted | c | 0.84 | 0.20 | 56 | 0.4 | 0.79 | Prompt attention |
| I9 | to allow your family and friends to take care of your personal needs, such as bringing you your favourite food or soap? | e | 0.92 | 0.17 | 79 | 2.4 | 0.75 | Access to social support (networks) (Social support) |
| I10 | <u>to practice religious or traditional observances</u> if you wanted to? | e | 0.95 | 0.14 | 89 | 11.7 | 0.62 | |
| I11 | allowed you to interact with family, friends and to continue your social and or religious customs during your stay | d | 0.83 | 0.16 | 45 | 1.6 | 0.65 | |

^a A before the item number refers to the ambulatory section of the questionnaire. Inpatient items were similarly worded. Inpatient exceptions are shown at the bottom and marked I ; ^b Shaded questions are adapted from CAHPS v2. Underlines refer to key phrases tested by translation and back-translation; ^c always(4) usually (3), sometimes (2), never (1); ^d very good(5), good(4), moderate(3), bad(2), very bad(1); ^e no problem(5), mild...(4), moderate...(3), severe...(2), extreme problem (1)

Internal consistency reliability estimates by socio-demographic breakdowns, shown in Table 2.2, were similar in most domains. The prompt attention domain was an exception: items showed higher reliability in more versus less educated populations in high human development countries (0.66 versus 0.60). Similar exceptions were noted for social support between males and females (0.71 versus 0.65), and younger and older groups (0.68 versus 0.63). No differences emerged in more refined age and education groups.

Temporal reliability

Across the 10 retest sites, average Kappa statistics ranged from 0.54 to 0.66 for ambulatory items, and from 0.59 to 0.79 for inpatient items (see Table 2.2). Of the ambulatory items, autonomy items had the highest reliability (K=0.66). Kappa statistics were higher for the same item for inpatient respondents ($p < 0.01$). Though Kappa statistics were generally adequate, there was

Table 2.2 Feasibility and reliability statistics for the responsiveness performance questions

| Description of sample statistics | Pooled Study dataset | | Country Study datasets | |
|---|----------------------|-----------|------------------------|-----------|
| | Ambulatory | Inpatient | Ambulatory | Inpatient |
| Observation points | | | | |
| All (face-to-face) surveys in the Study (n) | 50,876 | 7,964 | 41 | 41 |
| Low HDI Countries (n) | 36,500 | 5,306 | 24 | 24 |
| High HDI Countries (n) | 14,376 | 2,658 | 17 | 17 |
| Observation points for retests (9/10 were less developed countries) | 2,854 | 417 | 10 | 10 |
| Items | | | | |
| Number of items | 22 | 11 | 22 | 11 |
| Number of items with missing averages > 3% | 3 | 2 | 4 | 3 |
| Average item missing rate | 1.9% | 1.7% | 1.9% | 1.7% |
| Minimum average item missing rate | 0.2% | 0.1% | 0.2% | 0.1% |
| Maximum average item missing rate | 11.4% | 7.9% | 16.2% | 11.9% |
| Maximum country-item missing rate (41 countries) | n/a | n/a | 40.6% | 56.8% |
| Item ceiling effects | | | | |
| Number of times, last category >50% | 10 | 3 | 11 | 3 |
| Percent of times, last category >50% | 45% | 27% | 50% | 27% |
| Reliability: internal consistency- inter-item correlation coefficient | | | | |
| Prompt Attention | 0.50 | 0.69 | 0.53 | 0.66 |
| Dignity | 0.57 | n/a | 0.59 | n/a |
| Communication | 0.65 | n/a | 0.50 | n/a |
| Autonomy | 0.61 | n/a | 0.52 | n/a |
| Confidentiality | 0.62 | n/a | 0.54 | n/a |
| Choice of Care Provider | 0.59 | n/a | 0.54 | n/a |
| Quality of Basic Amenities | 0.82 | n/a | 0.74 | n/a |
| Social support | n/a | 0.40 | n/a | 0.39 |
| Reliability: domain sub-scales- alpha coefficients | | | | |
| Prompt Attention | 0.65 | 0.82 | 0.62 | 0.79 |
| Dignity | 0.84 | n/a | 0.81 | n/a |
| Communication | 0.88 | n/a | 0.87 | n/a |
| Autonomy | 0.82 | n/a | 0.78 | n/a |
| Confidentiality | 0.83 | n/a | 0.76 | n/a |
| Choice of Care Provider | 0.82 | n/a | 0.78 | n/a |
| Quality of Basic Amenities | 0.92 | n/a | 0.89 | n/a |
| Social support | n/a | 0.67 | n/a | 0.65 |

Table 2.2 Feasibility and reliability statistics for the responsiveness performance questions (continued)

| Description of sample statistics | Pooled Study dataset | | Country Study datasets | |
|---|----------------------|-----------|------------------------|-----------|
| | Ambulatory | Inpatient | Ambulatory | Inpatient |
| Reliability: internal consistency - alpha coefficients | | | | |
| Total or average: 1 scale for all items | 0.93 | 0.89 | 0.91 | 0.88 |
| Range country-item Alpha coefficient (41 countries, 2 scales) | n/a | n/a | 0.86-0.97 | 0.77-0.94 |
| Reliability: inter-temporal | | | | |
| Average Kappa statistic | 0.63 | 0.68 | 0.61 | 0.66 |
| Minimum item Kappa statistic | 0.58 | 0.59 | 0.54 | 0.59 |
| Maximum item Kappa statistic | 0.69 | 0.75 | 0.66 | 0.75 |
| Minimum country-item Kappa statistic (10 countries) | n/a | n/a | -0.09 | 0.0 |
| Maximum country-item Kappa statistic (10 countries) | n/a | n/a | 0.97 | 1.0 |

evidence of heterogeneity across retest sites (see Appendix 2.5). Six (of 10) retest sites, had Kappa statistics greater than 0.50 for all items. In Iran, statistics ranged between 0.40 and 0.60; in Georgia, between 0.30 and 0.55; in Columbia, between 0.24 and 0.50; and in Nigeria, between 0.20 and 0.40 except for communication and confidentiality items ($K < 0.2$).

Two-by-two comparisons for different socio-demographic groups showed significantly lower Kappa statistics in female, less educated, and less healthy populations ($p < 0.01$) than in the comparison groups. Correlations of Kappa statistics and age groups were modest ($r = 0.40$) and negative in 17 out of 33 cases. An examination of Kappa statistics by level of education confirmed that responses from less education populations were less reliable (correlation coefficients were positive in 30 of 33 cases).

Construct validity

Four factors in the ML factor analysis had eigenvalues > 1 , explaining 82% of the variance. The CAHPS communication items: listening (1.0), explaining (0.70), and time to ask questions (0.65), had the highest factor loadings on the unrotated general factor. Other items had loadings of 0.4 or more except "getting wanted care soon", "getting a provider you were happy with", and "using a provider other than your usual one". Other important factors were basic amenities, confidentiality, dignity and choice. The factor solution for developed countries contained 5 factors: a general factor, prompt attention-autonomy, basic amenities, communication, and confidentiality.

Less developed countries had 3 factors in the solution: a general factor, basic amenities, and choice.

Oblique promax rotated factor patterns are shown in Table 2.4. Item loadings of 0.4 or greater are bolded and underlined. Items expected to form part of a domain but with loadings <0.4 are only underlined. The rotated solutions confirmed the hypothesized domain taxonomy with few exceptions. The item on getting care as soon as you wanted (access) did not load on the same factor as getting prompt attention at health services (waiting time) for high human development countries. Also in high human development countries, the items for dignity tended to load on multiple factors. Not shown here are results obtained for sex, education, and health stratifications, which had similar rotated factor patterns. Correlations between factors ranged between 0.26 to 0.70 (average $r=0.45$). The factors with the highest correlation were confidentiality and communication (0.70). The lowest correlation was observed between the choice and respectful greetings factors ($r=0.26$).

Development, household income and age-based construct validity

Comparing more and less developed countries, we found average responsiveness performance to be higher in more developed settings for both ambulatory (0.84 versus 0.81, t-test, $p<0.07$) and inpatient (0.88 versus 0.86, $p<0.08$) respondents. Correlations between the HDI rank (the higher rank representing less development) and country-level responsiveness scores, overall for inpatient and ambulatory services and by domain, were generally negative (ambulatory average, -0.23 , $p=0.17$, inpatient average, -0.16 , $p=0.33$), and strongest for ambulatory dignity (-0.39 , $p=0.01$), confidentiality (-0.40 , $p=0.001$), and inpatient social support (-0.57 , $p=0.002$).

Income and responsiveness score correlations were positive in 24 (of 29) countries (average Gamma coefficient: 0.08, range: -0.13 to 0.30). Between 25% and 33% of correlations were statistically significant (Chi-square $p=0.10$) in both development contexts.

With increasing age, the proportion of respondents with higher responsiveness increased. Associations across age-groups within countries, measured by Gammas, were positive in 23 (of 24) cases in more developed countries (0.18, -0.09 to 0.49), in 11 (of 17) countries in less developed countries (0.05, -0.08 to 0.17), and significant ($p<0.10$) in about two-thirds of cases in each.

Table 2.3 Alpha coefficients for responsiveness performance questions, by domain, population group, and level of human development

| Domains | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice of (care) provider | Quality of basic amenities | Social support (hospital) | Prompt attention (hospital) |
|---|------------------|----------|---------------|----------|-----------------|---------------------------|----------------------------|---------------------------|-----------------------------|
| Populations in High Human Development Countries | n=14,376 | n=14,376 | n=14,376 | n=10,719 | n=14,376 | n=14,376 | n=14,376 | n=2,658 | n=2,658 |
| Female | 0.65 | 0.82 | 0.88 | 0.77 | 0.73 | 0.79 | 0.89 | 0.67 | 0.77 |
| Male | 0.65 | 0.81 | 0.88 | 0.76 | 0.74 | 0.79 | 0.88 | 0.65 | 0.79 |
| Younger | 0.64 | 0.81 | 0.87 | 0.78 | 0.74 | 0.79 | 0.88 | 0.65 | 0.78 |
| Older | 0.66 | 0.83 | 0.88 | 0.75 | 0.72 | 0.79 | 0.89 | 0.69 | 0.77 |
| Less educated | 0.60 | 0.80 | 0.86 | 0.78 | 0.75 | 0.79 | 0.90 | 0.65 | 0.75 |
| More educated | 0.67 | 0.82 | 0.88 | 0.76 | 0.73 | 0.79 | 0.88 | 0.66 | 0.80 |
| Healthy | 0.63 | 0.80 | 0.86 | 0.76 | 0.70 | 0.76 | 0.88 | 0.66 | 0.78 |
| Less healthy | 0.65 | 0.82 | 0.89 | 0.77 | 0.76 | 0.81 | 0.88 | 0.64 | 0.78 |
| Total | 0.65 | 0.81 | 0.88 | 0.77 | 0.74 | 0.79 | 0.88 | 0.66 | 0.78 |
| Populations in Low Human Development Countries | n=36,500 | n=36,500 | n=36,500 | n=25,704 | n=36,500 | n=36,500 | n=36,365 | n=5,306 | n=5,306 |
| Female | 0.67 | 0.85 | 0.89 | 0.83 | 0.85 | 0.83 | 0.94 | 0.65 | 0.84 |
| Male | 0.68 | 0.86 | 0.89 | 0.84 | 0.85 | 0.84 | 0.94 | 0.71 | 0.82 |
| Younger | 0.67 | 0.85 | 0.89 | 0.84 | 0.85 | 0.83 | 0.94 | 0.68 | 0.83 |
| Older | 0.67 | 0.85 | 0.89 | 0.84 | 0.85 | 0.85 | 0.95 | 0.63 | 0.82 |
| Less educated | 0.66 | 0.85 | 0.89 | 0.84 | 0.85 | 0.83 | 0.94 | 0.67 | 0.84 |
| More educated | 0.67 | 0.86 | 0.89 | 0.84 | 0.85 | 0.84 | 0.94 | 0.68 | 0.82 |
| Healthy | 0.66 | 0.85 | 0.88 | 0.84 | 0.85 | 0.83 | 0.94 | 0.68 | 0.81 |
| Less healthy | 0.67 | 0.86 | 0.89 | 0.83 | 0.85 | 0.84 | 0.94 | 0.67 | 0.84 |
| Total | 0.67 | 0.86 | 0.89 | 0.84 | 0.85 | 0.83 | 0.94 | 0.67 | 0.83 |

^a Following the skip pattern the number of respondents completing all 3 autonomy items was reduced

Table 2.4 Promax rotated factor solution for the ambulatory responsiveness performance questions

| Domain/ Question | High Human Development countries (n=14,376) | | | | | | | | | | | | Low Human Development countries (n=36,500) | | | | | | | | | | |
|-----------------------|---|-------------|-------------|-------------|-------------|-------|-------------|-------|-------------|-------------|--------------|-------|--|-------|-------------|-------|-------------|-------|--------------|-------------|-------|-------|----------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | U ^a | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | U ^a |
| Prompt Attention | A1 | -0.00 | <u>0.07</u> | 0.00 | -0.00 | 0.01 | -0.01 | 0.00 | -0.00 | -0.00 | 0.02 | 0.61 | 0.57 | -0.04 | 0.02 | 0.07 | 0.03 | -0.04 | 0.01 | <u>0.70</u> | -0.00 | 0.09 | 0.51 |
| | A2 | 0.00 | <u>0.98</u> | 0.01 | -0.00 | 0.01 | 0.00 | 0.00 | -0.00 | -0.01 | 0.00 | 0.01 | 0.00 | 0.13 | -0.05 | -0.05 | -0.01 | 0.04 | 0.03 | <u>0.60</u> | 0.01 | -0.18 | 0.38 |
| Dignity | A3 | <u>0.21</u> | -0.01 | 0.01 | -0.00 | -0.01 | 0.06 | -0.01 | -0.01 | <u>0.20</u> | -0.08 | 0.01 | 0.03 | 0.38 | -0.02 | -0.00 | 0.07 | -0.01 | 0.03 | <u>0.84</u> | 0.01 | 0.02 | -0.02 |
| | A4 | <u>1.00</u> | 0.00 | -0.00 | 0.00 | 0.01 | 0.02 | 0.00 | 0.00 | -0.02 | 0.01 | -0.00 | -0.01 | 0.00 | 0.01 | -0.02 | 0.04 | -0.00 | 0.02 | <u>0.92</u> | 0.00 | -0.02 | -0.02 |
| | A5 | 0.02 | -0.00 | 0.01 | 0.00 | 0.00 | 0.05 | 0.02 | 0.02 | <u>0.71</u> | -0.04 | 0.01 | -0.01 | 0.39 | <u>0.22</u> | -0.01 | 0.10 | 0.03 | -0.28 | <u>0.30</u> | -0.02 | 0.02 | 0.13 |
| | A6 | <u>0.04</u> | 0.06 | -0.02 | 0.01 | 0.04 | -0.02 | 0.02 | 0.01 | <u>0.13</u> | <u>-0.64</u> | -0.02 | -0.01 | 0.33 | <u>0.89</u> | -0.03 | 0.02 | 0.02 | 0.02 | 0.04 | 0.02 | -0.00 | -0.01 |
| Communication | A7 | -0.05 | 0.01 | 0.00 | -0.01 | 0.01 | <u>0.50</u> | -0.02 | 0.02 | -0.05 | -0.03 | 0.02 | -0.04 | 0.34 | -0.00 | -0.02 | <u>0.64</u> | 0.01 | -0.03 | 0.07 | 0.00 | 0.34 | -0.01 |
| | A8 | 0.03 | 0.01 | 0.02 | -0.01 | 0.00 | <u>0.77</u> | 0.00 | 0.02 | 0.01 | 0.06 | 0.00 | -0.00 | 0.32 | -0.02 | -0.01 | <u>0.79</u> | 0.02 | -0.01 | 0.01 | 0.00 | 0.10 | -0.03 |
| | A9 | 0.03 | 0.00 | -0.00 | 0.01 | 0.01 | <u>0.83</u> | 0.02 | -0.01 | 0.06 | -0.01 | 0.01 | 0.01 | 0.24 | -0.00 | -0.02 | <u>0.87</u> | 0.01 | -0.03 | 0.02 | -0.00 | -0.16 | -0.03 |
| | A10 | -0.01 | 0.00 | 0.04 | 0.01 | 0.01 | <u>0.38</u> | 0.01 | -0.02 | -0.08 | <u>-0.59</u> | -0.03 | 0.03 | 0.28 | 0.18 | -0.02 | <u>0.49</u> | -0.03 | 0.05 | -0.00 | -0.00 | 0.00 | -0.39 |
| Autonomy ^b | A11 | 0.00 | 0.01 | <u>0.98</u> | 0.00 | 0.01 | 0.03 | 0.01 | 0.00 | 0.01 | 0.01 | -0.00 | -0.00 | 0.00 | -0.00 | -0.04 | 0.17 | 0.02 | -0.09 | 0.06 | 0.06 | -0.00 | <u>-0.45</u> |
| Confidentiality | A14 | 0.00 | 0.00 | 0.00 | <u>0.99</u> | 0.00 | 0.00 | -0.01 | 0.01 | -0.00 | -0.02 | 0.01 | -0.00 | 0.00 | -0.02 | -0.04 | 0.07 | 0.01 | <u>-0.74</u> | 0.00 | 0.03 | -0.03 | -0.01 |
| | A15 | -0.00 | -0.00 | -0.02 | 0.10 | 0.00 | 0.03 | 0.03 | <u>0.78</u> | 0.04 | 0.11 | -0.10 | 0.03 | 0.36 | -0.04 | 0.01 | 0.02 | 0.00 | <u>-0.87</u> | -0.01 | 0.00 | 0.02 | -0.02 |
| | A16 | 0.02 | 0.01 | 0.05 | -0.02 | 0.02 | -0.02 | -0.04 | <u>0.61</u> | -0.03 | -0.20 | 0.17 | -0.04 | 0.37 | 0.06 | -0.02 | -0.07 | -0.03 | <u>-0.63</u> | -0.03 | -0.03 | 0.00 | <u>-0.44</u> |

Table 2.4 Promax rotated factor solution for the ambulatory responsiveness performance questions (continued)

| Domain/ Question | High Human Development countries (n=14,376) | | | | | | | | | | | | Low Human Development countries (n=36,500) | | | | | | | | | U ^a | | |
|-------------------------------|---|-------|-------|-------|-------|-------------|-------|-------------|-------|-------|-------|-------------|--|------|-------|--------------|-------|-------------|-------|-------|-------|----------------|-------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | | |
| Choice of (care) provider | A17 | 0.01 | 0.00 | 0.02 | 0.01 | -0.01 | 0.02 | 0.68 | 0.03 | -0.05 | -0.06 | -0.01 | 0.04 | 0.39 | 0.00 | 0.01 | 0.02 | 0.80 | -0.01 | 0.02 | -0.02 | -0.01 | -0.01 | 0.34 |
| | A18 | 0.00 | 0.00 | -0.00 | -0.01 | 0.02 | 0.00 | 0.88 | -0.02 | 0.04 | 0.02 | 0.03 | -0.03 | 0.26 | 0.01 | -0.00 | 0.00 | 0.87 | 0.03 | -0.03 | 0.01 | 0.01 | 0.04 | 0.29 |
| | A19 | -0.00 | 0.01 | 0.02 | 0.02 | 0.00 | 0.02 | 0.18 | 0.01 | 0.01 | 0.04 | 0.69 | 0.02 | 0.31 | 0.00 | -0.05 | -0.01 | 0.52 | -0.05 | 0.02 | 0.01 | 0.00 | -0.34 | 0.34 |
| Quality of basic amenities | A20 | 0.01 | 0.01 | -0.00 | 0.02 | 0.83 | 0.00 | 0.01 | -0.03 | -0.02 | 0.05 | 0.03 | -0.01 | 0.30 | -0.01 | -0.87 | 0.00 | 0.02 | -0.00 | 0.01 | 0.02 | -0.00 | -0.01 | 0.21 |
| | A21 | -0.00 | -0.03 | 0.01 | -0.01 | 0.80 | -0.00 | -0.04 | 0.02 | 0.02 | -0.03 | 0.01 | 0.05 | 0.32 | 0.02 | -0.93 | 0.02 | -0.01 | -0.01 | -0.01 | -0.01 | 0.01 | 0.03 | 0.15 |
| | A22 | 0.00 | 0.02 | 0.00 | -0.01 | 0.90 | 0.01 | 0.04 | 0.01 | 0.01 | -0.03 | -0.04 | -0.03 | 0.20 | -0.00 | -0.92 | 0.00 | -0.00 | -0.00 | 0.01 | -0.01 | -0.01 | -0.02 | 0.13 |

^a Unique variance; ^b Note that 2 of the 3 autonomy questions formed part of a skip pattern and are excluded from this analysis

DISCUSSION

Responses from respondents from different countries, cultures, and health care systems displayed psychometric properties that compared well with other surveys measuring similar domains of quality of care. The article on the 19-item, ambulatory CAHPS surveys of quality of care in the US ($n=166,074$)¹², and the 5-country, 15-item, inpatient Picker survey ($n=62,925$)²⁴ provide illustrative comparisons. Inappropriate item missing rates for CAHPS ranged from 3% to 66% across items and plans compared with 0%-57% across items and countries for our study. CAHPS's confirmatory factor analysis had similar factor loadings (0.39-0.82). Alpha coefficients in the Picker and WHO studies had similar ranges (0.80 to 0.90).

Other instruments with comparable psychometric properties were identified in a meta-analysis of patient satisfaction surveys.²⁵ Dignity and communication were associated in our study and in another CAHPS studies.²⁶ Ceiling effects and conceptual equivalence properties were similar to those for health questions in the 11-country International Quality of Life Assessment (IQOLA) project.¹⁸

Temporal reliability was modest. Poor results in 4 of 10 sites may have arisen from poor translations, yet item equivalence for all 10 surveys was similar. Administration procedures or changes in the respondent's health service utilization are other explanations, but conclusive explanations cannot be provided without further testing. The observed lower temporal reliability for respondents with lower education was similar to observations on the measurement of self-reported health.²⁷ The observed lower reliability for females reflected education biases (females had 1.5 less years of education than males) and females' more complex health needs²⁸: preliminary exploration showed lower reliability even when education was held constant.

Judging from the factor solutions, the domains of responsiveness were understood by different populations as distinct, complementary aspects of quality. The stronger factorial loading of the autonomy item in the analysis for countries with more compared with less human development supported the view that users' desires to be involved in decision-making is related to general human development and education, rather than being purely a cultural artefact. Internal consistency results for most domain scales were good, particularly for basic amenities, communication, dignity, and autonomy, where missing rates were also low. The test of mean rankings showed preliminary evidence for translation equivalence across items, though it was notably weaker in low human development countries.

Although most items displayed considerable psychometric strengths, certain items need improving across all settings: most notably in the domains of confidentiality, choice, prompt at-

tention, and social support. The high missing rates for confidentiality items indicated that they covered issues beyond the knowledge of general populations when stated so generally. The choice item, "using a provider other than your usual one", had high inappropriate non-response rates and was probably confusing (see Picker surveys for promising alternatives for choice²⁹). As seen in other studies covering prompt attention, the items on waiting and access were not unidimensional.³⁰ Uneven alpha results for the social support 'religion' item may have reflected real differences in the relevance of religion in health care delivery settings across societies and could be omitted if irrelevant.

As is, the responsiveness item questions captured four out of the five areas of patient-physician relationships identified by Stewart.³¹ Exceptions were "positive affect, empathy, warmth and encouragement". Though possibly covered in dignity or communication domains, adding specific questions on empathy and care coordination would increase the coverage of women's health issues.²⁸ None of the above excludes shortening the questionnaire, which should be investigated to reduce costs.

This paper provides preliminary evidence that raw responses to the responsiveness questions reflect the realities of care quality. For example, more developed countries reported better responsiveness. However, associations with income were weak. Also, the observed positive correlations of responsiveness with age were similar to Campbell et al.'s³² findings from the 1970 study of the quality of American life, which they attributed to an objective, improved fit of people with their circumstances over time, *in addition* to declining expectations. The lingering impact of expectations on outcome measurement hinders interpretation but may be mitigated if expectations are measured and used to adjust raw responses. Measurement may also benefit from techniques allowing patients opportunities for constructive criticism and to explain context.³³

A general weakness of the MCS Study was its low survey response rates (though we found no similar studies with as many challenges in the form of low development settings).³⁴ Although this made our analysis of responsiveness vulnerable to bias arising from systematic differences in non-respondents associated with quality of care evaluations¹⁵, given the convergence of our principal findings across all the subgroups analysed, and given that certain subgroups were likely to describe dominant socio-demographic characteristics of non-respondents (e.g., less educated), we concluded that our key results were robust.

In summary, the lessons provided here from 41 countries build on existing work and provide useful insights on how to conduct and improve upon measurement of non-clinical processes of care for population groups in different settings.

REFERENCES

1. Kelley E, Hurst J. Health care quality indicators project conceptual framework paper. Paris: OECD; 2006.
2. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
3. Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bull World Health Organ.* 2000;78(6):717–31.
4. Thompson AGH, Sunol R. Expectations as determinants of patient satisfaction. *Health Expect.* 1995;2:93–104.
5. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med.* 1997;45(12):1829–43.
6. Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. *Soc Sci Med.* 1998;47(10):1573–88.
7. Ware JE, Hays RD. Methods for measuring patient satisfaction with specific medical encounters. *Med Care.* 1998;26(4):393–402.
8. Penchansky R, Thomas JW. The concept of access. *Med Care.* 1981;19(2):127–40.
9. de Silva A. A framework for measuring responsiveness. Global programme on evidence for health policy discussion paper no. 32. Geneva: World Health Organization; 2000.
10. Zastowny TR, Stratmann WC, Adams EJ, Fox ML. Patient satisfaction and experience with health services and quality of care. *Qual Manag Health.* 1995;C(3):50–61.
11. Murray CJL, Kawabata K, Valentine NB. People's experience versus people's expectations. *Health Affair.* 2001;20(3):41–3.
12. Hargraves JL, Hays RD, Cleary PD. Psychometric properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 adult core survey. *Health Serv Res.* 2003;38(6 Pt 1):1509–27.
13. Üstün TB, Chatterji S, Villanueva M, Bendib L, Celik C, Sadana R, Valentine NB, Ortiz J, Tandon A, Salomon J, Cao Y, Wan Jun X, Özaltin E, Mathers C, Murray CJL. WHO Multi-Country Survey Study on Health and Responsiveness 2000–2001. Global programme on evidence for health policy. Discussion paper no. 37. Geneva: World Health Organization; 2001.
14. Schafer JL. NORM: multiple imputation of incomplete multivariate data under a normal model [statistical software]. University Park: Penn. State University, Department of Statistics; 2001.
15. Elliott MN, Edwards C, Hambarsoomians K, Angeles J, Hays RD. Patterns of unit and item nonresponse in the CAHPS Hospital Survey. *Health Serv Res.* 2005;40(6p2):2096–119.
16. DeVellis RF. Scale development. Theory and applications. London: Sage; 1991.
17. United Nations Development Agency (UNDP). Human development report. New York: UNDP; 2001.
18. Gandek B, Ware JE, Jr., Aaronson NK, Alonso J, Apolone G, Bjorner J, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. International Quality of Life Assessment. *J Clin Epidemiol.* 1998;51(11):1149–58.

19. Cramer D, Howitt D. *The SAGE dictionary of statistics*. London: Sage; 2004.
20. Nunnally JCB, Bernstein IH. *Psychometric theory*. 3rd ed. New York: McGraw-Hill; 1994.
21. Kim J, Mueller CW. Factor analysis statistical methods and practical issues. In: Uslaner EM, editor. *Quantitative applications in the social sciences*. London: Sage; 1978.
22. Hall JA, Dornan MC. Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis. *Soc Sci Med*. 1990;30:811–8.
23. United Nations Educational, Scientific and Cultural Organization(UNESCO). Paris: UNESCO; 2002.
24. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Health Care*. 2002;14(5):353–8.
25. Sitzia J. How valid and reliable are patient satisfaction data? An analysis of 195 studies. *Int J Qual Health Care*. 1999;11(4):319–28.
26. Sofaer S, Crofton E, Goldstein C, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? *Health Serv Res*.40(6p2):2018–37.
27. Ware J. Scales for measuring general health perceptions. *Health Serv Res*. 1976 Winter; 396–415.
28. Hudson S, Weisman C, Anderson R, Camacho F. The development and validation of the primary care satisfaction survey for women. *Women Health Iss*. 2004 Mar-Apr;14(2):35–50.
29. Coulter A, Jenkinson C. European patients' views on the responsiveness of health systems and healthcare providers. *Eur J Public Health*. 2005;15(4):355–60.
30. Grogan S, Conner M, Norman P, Willits D, Porter I. Validation of a questionnaire measuring patient satisfaction with general practitioner services. *Qual Safe Health Care*. 2000;(9):210–5.
31. Stewart M. Reflections on the doctor–patient relationship from evidence and experience. *Brit J Gen Pract*. 2005;55(519):793–801.
32. Campbell A, Converse PE, Rodgers WL. *The quality of American life; perceptions evaluations and satisfactions*. New York: Russell Sage Foundation; 1976.
33. Edwards C, Staniszwesda S, Crichton N. Investigation of the ways in which patients' reports of their satisfaction with healthcare are constructed. *Sociol Health Ill*. 2004;26(2):159–83.
34. Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care*. 1998;10(4):311–7.

ENDNOTES

ⁱ High human development - Argentina, Bahrain, Belgium, Canada (telephone), Costa Rica, Croatia, Czech Republic, Estonia, Finland, France, Germany, Iceland, Ireland, Italy, Latvia, Luxembourg, Malta, Mexico (long survey - 9 modules), Portugal, Slovakia (long), Spain, Sweden, The Netherlands, United Arab Emirates. Low human development - Bulgaria, China (long), Colombia (long), Egypt (long), Georgia (long), India (long), Indonesia (long), Iran (long), Jordan, Morocco, Nigeria (long), Oman, Romania, Russian Federation, Syria (long), Turkey (long), Venezuela.

Appendix 2.1 Performance questions from the responsiveness module in the WHO MCS Study

| Ambulatory (A) or Inpatient(I) ^c | Question wording ^a | Response options (numeric coding in parentheses) | Domain |
|---|--|---|-----------------------|
| A1 | In the last 12 months, when you wanted care, how often did you get care <u>as soon as you wanted?</u> ^d | Always(4) usually (3), sometimes (2), never (1) | PROMPT ATTENTION (PA) |
| A2 | Overall, how would you rate your experience of <u>getting prompt attention at the health services</u> in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| A3 | In the last 12 months, when you sought care, how often did doctors ^b <u>treat you with respect?</u> ^e | always(4) usually (3), sometimes (2), never (1) | |
| A4 | In the last 12 months, when you sought care, how often did the office staff, such as receptionists or clerks there, <u>treat you with respect?</u> ^e | always(4) usually (3), sometimes (2), never (1) | DIGNITY (DIG) |
| A5 | In the last 12 months, how often were your <u>physical examinations and treatments done in a way that your privacy was respected?</u> | always(4) usually (3), sometimes (2), never (1) | |
| A6 | Overall, how would you rate your experience of <u>getting treated with dignity</u> at the health services in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| A7 | In the last 12 months, how often did doctors ^b <u>listen carefully to you?</u> | always(4) usually (3), sometimes (2), never (1) | |
| A8 | In the last 12 months, how often did doctors ^b there, <u>explain things in a way you could understand?</u> | always(4) usually (3), sometimes (2), never (1) | COMMUNICATION (COM) |
| A9 | In the last 6 months, how often did doctors ^b give you time to <u>ask questions about your health problem or treatment?</u> | always(4) usually (3), sometimes (2), never (1) | |
| A10 | Overall, how would you rate your experience of <u>how well health care providers communicated with you</u> in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| A11 | In the last 12 months, how often did doctors ^b there <u>involve you as much as you wanted to be in deciding about the care..?</u> (preceded by a skip) | always(4) usually (3), sometimes (2), never (1) | |
| A12 | In the last 12 months, how often did doctors ^b there <u>ask your permission</u> before starting tests or treatment? | always(4) usually (3), sometimes (2), never (1) | AUTONOMY (AUT) |
| A13 | Overall, how would you rate your experience of getting involved in making decisions about your care or treatment as much as you wanted in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| A14 | In the last 12 months, how often were talks with your doctor ^b done privately so other people who you did not want to hear could not overhear what was said? | always(4) usually (3), sometimes (2), never (1) | |
| A15 | In the last 12 months, how often did your doctor ^b keep your personal information confidential? This means that anyone whom you did not want informed could not find out about your medical conditions. | always(4) usually (3), sometimes (2), never (1) | CONFIDENTIALITY (CON) |
| A16 | Overall, how would you rate your experience of <u>the way the health services kept information about you confidential</u> in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |

Appendix 2.1 continued

| Ambulatory (A) or Inpatient(I) ^c | Question wording ^a | Response options (numeric coding in parentheses) | Domain |
|---|--|--|----------------------------------|
| A17 | In the last 12 months, with the doctors ^b available to you how big a problem, if any, was it to <u>get to a health care provider you were happy with?</u> | no problem(5), mild..(4), moderate..(3), severe..(2), extreme problem..(1) | CHOICE OF PROVIDER (CH) |
| A18 | Over the last 12 months, how big a problem, if any, was it to <u>get to use other health care services other than the one you usually went to?</u> | no problem(5), mild..(4), moderate..(3), severe..(2), extreme problem..(1) | |
| A19 | Overall, how would you rate your experience of <u>being able to use a health care provider or service of your choice</u> over the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| A20 | Thinking about the places you visited for health care in the last 12 months, how would you rate the basic quality of the waiting room, for example, space, seating and fresh air? | very good(5), good(4), moderate(3), bad(2), very bad(1) | QUALITY OF BASIC AMENITIES (QBA) |
| A21 | Thinking about the places you visited for health care over the last 12 months, how would you rate the cleanliness of the place? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| A22 | Overall, how would you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanliness of the health services you visited in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| I1 | When you were in the hospital, how often did you get attention from doctors as quickly as you wanted? | always(4) usually (3), sometimes (2), never (1) | PA |
| I2 | Overall, how would you rate your experience of getting prompt attention at the hospital in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |
| I3 | Overall, how would you rate your experience of getting treated with dignity at the hospital in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | DIG |
| I4 | Overall, how would you rate your experience of how well health care providers communicated with you during your stay in the hospital in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | COM |
| I5 | Overall, how would you rate your experience of getting involved in making decisions about your care or treatment as much as you wanted when you were in hospital in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | AUT |
| I6 | Overall, how would you rate your experience of the way the hospital kept personal information about you confidential in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | CON |
| I7 | Overall, how would you rate your experience of being able to use a hospital of your choice over the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | CH |
| I8 | Overall, how would you rate <u>the overall quality of the surroundings</u> , for example, space, seating, fresh air, and cleanliness of the health services you visited in the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | QBA |

Appendix 2.1 continued

| Ambulatory (A) or Inpatient (I) ^c | Question wording ^a | Response options (numeric coding in parentheses) | Domain |
|--|---|--|-------------------------------------|
| 19 | In the last 12 months, when you stayed in hospital, how big a problem, if any, was it to get the hospital to allow your family and friends to take care of your personal needs, such as bringing you your favourite food or soap? | no problem(5), mild..(4), moderate..(3), severe..(2), extreme problem..(1) | ACCESS TO SOCIAL SUPPORT (NETWORKS) |
| 110 | During your stay in hospital, how big a problem, if any, was it to have the hospital allow you to <u>practice religious or traditional observances</u> if you wanted to? | no problem(5), mild..(4), moderate..(3), severe..(2), extreme problem..(1) | |
| 111 | Now, overall, how would you rate your experience of how the hospital allowed you to interact with family, friends and to continue your social and or religious customs during your stay over the last 12 months? | very good(5), good(4), moderate(3), bad(2), very bad(1) | |

^a Shaded questions are adapted from CAHPS v2; ^b All questions referring to “doctors”, were followed by the phrase “nurses or other health care providers”; ^c Ambulatory and inpatient questions were grouped as two different sections in the questionnaire; ^d CAHPS used the phrase “when you needed care right away” rather than “when you wanted care”; ^e CAHPS used the phrase “courtesy and respect” rather than just “respect”

Appendix 2.2 Sample descriptive statistics for the 41 surveys in the WHO MCS Study

| Survey ('long' refers to questionnaires containing 9 modules; the others contained 3 or 4 modules) | Response rate for attempted contacts (effective contacts) (%) | Ambulatory users of health services in previous 12 months | Hospital inpatient users in previous 12 months | Percentage female | Average age (years) | Average education level (years) | Good or very good health (%) |
|--|---|---|--|-------------------|---------------------|---------------------------------|------------------------------|
| <i>High Human Development</i> | | 14,376 | 2,658 | 53 | 45 | 12 | 56 |
| Argentina ^a | 36(.) | 478 | 87 | 59 | 45 | 10 | 67 |
| Bahrain ^a | 35(44) | 393 | 82 | 51 | 35 | 11 | 87 |
| Belgium ^a | 36(48) | 618 | 134 | 57 | 46 | 13 | 62 |
| Canada (telephone) | 11(11) | 164 | 39 | 40 | 48 | 14 | 68 |
| Costa Rica | 37(67) | 498 | 78 | 41 | 40 | 8 | 58 |
| Croatia | .(68) | 845 | 161 | 67 | 50 | 10 | 39 |
| Czech Republic ^a | 55(60) | 717 | 148 | 35 | 46 | 14 | 51 |
| Estonia | 50(71) | 725 | 152 | 61 | 47 | 10 | 33 |
| Finland | 21(52) | 722 | 154 | 58 | 48 | 14 | 51 |
| France ^a | 42(77) | 652 | 131 | 57 | 44 | 18 | 71 |
| Germany | 67(80) | 698 | 96 | 53 | 50 | 13 | 59 |
| Iceland | .(53) | 302 | 42 | 60 | 39 | 20 | 72 |
| Ireland ^a | 17(39) | 337 | 89 | 55 | 47 | 12 | 66 |

Appendix 2.2 continued

| Survey ('long' refers to questionnaires containing 9 modules; the others contained 3 or 4 modules) | Response rate for attempted contacts (effective contacts) (%) | Ambulatory users of health services in previous 12 months | Hospital inpatient users in previous 12 months | Percentage female | Average age (years) | Average education level (years) | Good or very good health (%) |
|--|---|---|--|-------------------|---------------------|---------------------------------|------------------------------|
| Italy ^a | 36(61) | 451 | 69 | 55 | 48 | 12 | 46 |
| Latvia | 54(72) | 445 | 119 | 44 | 49 | 12 | 25 |
| Luxembourg ^a (telephone) | (.55) | 512 | 96 | 57 | 45 | 14 | 69 |
| Malta | 48(59) | 322 | 51 | 53 | 48 | 12 | 54 |
| Mexico ^a (long) | (.96) | 1,833 | 332 | 33 | 45 | 9 | 55 |
| Portugal ^a | 37(61) | 534 | 95 | 62 | 49 | 8 | 26 |
| Slovakia (long) | 39(84) | 796 | 140 | 60 | 44 | 12 | 48 |
| Spain | 19(75) | 619 | 85 | 58 | 48 | 11 | 49 |
| Sweden | 46(53) | 564 | 107 | 57 | 49 | 10 | 59 |
| The Netherlands | 55(59) | 686 | 84 | 58 | 45 | 14 | 59 |
| United Arab Emirates | 71(76) | 465 | 87 | 52 | 35 | 12 | 76 |
| <i>Low Human Development</i> | | 36,500 | 5,306 | 57 | 40 | 9 | 55 |
| Bulgaria | 69(88) | 583 | 82 | 55 | 47 | 14 | 43 |
| China ^b (long) | (.99) | 4,441 | 783 | 49 | 42 | 9 | 62 |
| Colombia (long) | 72(84) | 3,840 | 532 | 72 | 41 | 8 | 56 |
| Egypt (long) | (.99) | 2,610 | 248 | 61 | 42 | 7 | 60 |
| Georgia (long) | 84(93) | 2,172 | 357 | 63 | 47 | 12 | 18 |
| India ^b (long) | (.98) | 3,172 | 434 | 56 | 41 | 4 | 58 |
| Indonesia (long) | (.99) | 3,971 | 313 | 61 | 42 | 7 | 64 |
| Iran (long) | not available | 7,101 | 884 | 59 | 38 | 6 | 54 |
| Jordan | 74(83) | 373 | 97 | 57 | 36 | 10 | 73 |
| Morocco ^a | (.69) | 434 | 55 | 55 | 37 | 8 | 48 |
| Nigeria ^{a, b} (long) | 45(98) | 1,005 | 170 | 63 | 37 | 8 | 71 |
| Oman | 67(79) | 519 | 86 | 46 | 33 | 11 | 76 |
| Romania ^a | 39(52) | 538 | 170 | 56 | 46 | 13 | 38 |
| Russian Federation | 25(10) | 947 | 208 | 59 | 43 | 15 | 29 |
| Syria (long) | not available | 3,149 | 701 | 60 | 40 | 7 | 53 |
| Turkey ^a (long) | (.90) | 1,381 | 130 | 54 | 35 | 10 | 65 |
| Venezuela | (.66) | 264 | 56 | 40 | 35 | 11 | 67 |
| Total or average | 46(70) | 50,876 | 7,964 | 57 | 40 | 9 | 53 |

^a Income variable missing rate>15%; ^b Sub-national samples: China (Shandon, Henan and Gansu);

India (Andhra Pradesh); Nigeria (Oyo State)

Appendix 2.3 Likert-scaled responsiveness inpatient items and item properties

| Item | Item handles | Response options | Domain | Standardized mean (0-1) | Standard deviation | Ceiling effect (%) | Missing rate (%) | Kappa statistic (0-1) |
|------|---|------------------|--|-------------------------|--------------------|--------------------|------------------|-----------------------|
| 11 | getting attention from doctors as quickly as you wanted | a | PROMPT ATTENTION | 0.84 | 0.20 | 56 | 0.4 | 0.79 |
| 12 | getting prompt attention at the hospital in the last 12 months | b | | 0.80 | 0.17 | 33 | 0.1 | 0.68 |
| 13 | getting treated with dignity at the hospital in the last 12 months? | b | DIG-NITY | 0.82 | 0.16 | 36 | 0.1 | 0.71 |
| 14 | how well health care providers communicated with you during your stay in the hospital | b | COMMUN- CATION | 0.80 | 0.17 | 33 | 0.2 | 0.71 |
| 15 | getting involved in making decisions about your care or treatment as much as you wanted when you were in hospital | b | AUTON- OMY | 0.76 | 0.19 | 26 | 1.6 | 0.66 |
| 16 | the way the hospital kept personal information about you confidential | b | CON- FIDEN- TIALTY | 0.82 | 0.16 | 39 | 7.1 | 0.60 |
| 17 | being able to use a hospital of your choice | b | CHOICE OF CARE PROVIDER | 0.77 | 0.20 | 31 | 5.4 | 0.63 |
| 18 | <u>the overall quality of the surroundings</u> , for example, space, seating, fresh air, and cleanliness of the health services | b | QUALITY OF BASIC AMENITIES | 0.78 | 0.18 | 31 | 0.1 | 0.64 |
| 19 | to allow your family and friends to take care of your personal needs, such as bringing you your favourite food or soap? | c | ACCESS TO SOCIAL SUPPORT (NETWORKS) (SOCIAL SUPPORT) | 0.92 | 0.17 | 79 | 2.4 | 0.75 |
| 110 | <u>to practice religious or traditional observances</u> if you wanted to? | c | | 0.95 | 0.14 | 89 | 11.7 | 0.62 |
| 111 | allowed you to interact with family, friends and to continue your social and or religious customs during your stay | b | | 0.83 | 0.16 | 45 | 1.6 | 0.65 |

^a always(4) usually (3), sometimes (2), never (1); ^b very good(5), good(4), moderate(3), bad(2), very bad (1); ^c no problem(5), mild..(4), moderate..(3), severe..(2), extreme problem (1)

Appendix 2.4(a) Internal consistency reliability in more developed countries (n=24)

| Country study | All questions (ambulatory) | All questions (hospital) | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice of care provider | Quality of basic amenities | Social support | Prompt attention (hospital) |
|---|----------------------------|--------------------------|------------------|---------|---------------|----------|-----------------|-------------------------|----------------------------|----------------|-----------------------------|
| Alpha coefficients (for n see Appendix 2.2) | | | | | | | | | | | |
| Argentina | 0.88 | 0.81 | 0.43 | 0.70 | 0.83 | 0.71 | 0.71 | 0.56 | 0.87 | 0.50 | 0.72 |
| Bahrain | 0.88 | 0.84 | 0.49 | 0.80 | 0.84 | 0.91 | 0.86 | 0.69 | 0.95 | 0.73 | 0.80 |
| Belgium | 0.91 | 0.86 | 0.45 | 0.76 | 0.89 | 0.76 | 0.59 | 0.73 | 0.87 | 0.39 | 0.86 |
| Canada | 0.91 | 0.84 | 0.66 | 0.69 | 0.88 | 0.77 | 0.74 | 0.80 | 0.87 | 0.59 | 0.81 |
| Costa Rica | 0.89 | 0.88 | 0.58 | 0.74 | 0.84 | 0.74 | 0.74 | 0.69 | 0.87 | 0.62 | 0.84 |
| Croatia | 0.93 | 0.92 | 0.56 | 0.82 | 0.88 | 0.82 | 0.75 | 0.80 | 0.90 | 0.83 | 0.82 |
| Czech Republic | 0.94 | 0.90 | 0.68 | 0.86 | 0.89 | 0.81 | 0.76 | 0.89 | 0.85 | 0.66 | 0.81 |
| Estonia | 0.89 | 0.92 | 0.69 | 0.78 | 0.83 | 0.72 | 0.56 | 0.81 | 0.86 | 0.79 | 0.85 |
| Finland | 0.90 | 0.81 | 0.69 | 0.82 | 0.83 | 0.79 | 0.71 | 0.76 | 0.89 | 0.63 | 0.80 |
| France | 0.90 | 0.90 | 0.50 | 0.81 | 0.86 | 0.75 | 0.71 | 0.70 | 0.88 | 0.61 | 0.78 |
| Germany | 0.92 | 0.90 | 0.67 | 0.83 | 0.86 | 0.64 | 0.73 | 0.72 | 0.84 | 0.72 | 0.73 |
| Iceland | 0.90 | 0.85 | 0.70 | 0.74 | 0.88 | 0.79 | 0.79 | 0.79 | 0.85 | 0.83 | 0.71 |
| Ireland | 0.93 | 0.90 | 0.72 | 0.78 | 0.91 | 0.77 | 0.56 | 0.84 | 0.91 | 0.64 | 0.83 |
| Italy | 0.93 | 0.91 | 0.56 | 0.81 | 0.89 | 0.69 | 0.75 | 0.81 | 0.89 | 0.70 | 0.91 |
| Latvia | 0.88 | 0.81 | 0.55 | 0.66 | 0.80 | 0.75 | 0.71 | 0.75 | 0.85 | 0.29 | 0.65 |
| Luxembourg | 0.86 | 0.84 | 0.47 | 0.73 | 0.85 | 0.70 | 0.33 | 0.64 | 0.79 | 0.53 | 0.78 |
| Malta | 0.90 | 0.87 | 0.63 | 0.81 | 0.88 | 0.78 | 0.66 | 0.72 | 0.89 | 0.59 | 0.77 |
| Mexico | 0.92 | 0.89 | 0.71 | 0.85 | 0.91 | 0.74 | 0.81 | 0.78 | 0.92 | 0.68 | 0.72 |
| Portugal | 0.91 | 0.90 | 0.40 | 0.84 | 0.89 | 0.71 | 0.72 | 0.81 | 0.89 | 0.58 | 0.79 |
| Slovakia | 0.94 | 0.91 | 0.70 | 0.89 | 0.90 | 0.84 | 0.85 | 0.91 | 0.85 | 0.80 | 0.85 |

Appendix 2.4(a) Internal consistency reliability in more developed countries (n=24) (continued)

| Country study | All questions (ambulatory) | All questions (hospital) | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice of care provider | Quality of basic amenities | Social support | Prompt attention (hospital) |
|----------------------|----------------------------|--------------------------|------------------|---------|---------------|----------|-----------------|-------------------------|----------------------------|----------------|-----------------------------|
| Spain | 0.90 | 0.89 | 0.68 | 0.82 | 0.88 | 0.80 | 0.78 | 0.67 | 0.86 | 0.56 | 0.85 |
| Sweden | 0.92 | 0.83 | 0.78 | 0.90 | 0.89 | 0.72 | 0.61 | 0.84 | 0.86 | 0.64 | 0.59 |
| The Netherlands | 0.89 | 0.77 | 0.71 | 0.76 | 0.84 | 0.70 | 0.62 | 0.74 | 0.82 | 0.41 | 0.54 |
| United Arab Emirates | 0.93 | 0.92 | 0.46 | 0.82 | 0.87 | 0.81 | 0.78 | 0.79 | 0.90 | 0.75 | 0.85 |
| Average | 0.91 | 0.87 | 0.60 | 0.79 | 0.87 | 0.76 | 0.70 | 0.76 | 0.87 | 0.63 | 0.78 |
| Standard deviation | 0.02 | 0.04 | 0.11 | 0.06 | 0.03 | 0.06 | 0.11 | 0.08 | 0.03 | 0.14 | 0.09 |

Appendix 2.4(b) Internal consistency reliability in less developed countries (n=17)

| Country study | All questions (ambulatory) | All questions (hospital) | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice of care provider | Quality of basic amenities | Social support | Prompt attention (hospital) |
|---|----------------------------|--------------------------|------------------|---------|---------------|----------|-----------------|-------------------------|----------------------------|----------------|-----------------------------|
| Alpha coefficients (for n see Appendix 2.2) | | | | | | | | | | | |
| Bulgaria | 0.95 | 0.89 | 0.73 | 0.88 | 0.90 | 0.85 | 0.86 | 0.86 | 0.94 | 0.76 | 0.84 |
| China | 0.92 | 0.87 | 0.65 | 0.86 | 0.86 | 0.86 | 0.83 | 0.80 | 0.92 | 0.59 | 0.79 |
| Colombia | 0.91 | 0.89 | 0.66 | 0.81 | 0.88 | 0.83 | 0.78 | 0.86 | 0.94 | 0.76 | 0.85 |
| Egypt | 0.97 | 0.93 | 0.81 | 0.93 | 0.95 | 0.89 | 0.91 | 0.88 | 0.97 | 0.63 | 0.91 |
| Georgia | 0.92 | 0.92 | 0.67 | 0.81 | 0.88 | 0.86 | 0.88 | 0.90 | 0.97 | 0.80 | 0.81 |
| India | 0.90 | 0.87 | 0.56 | 0.82 | 0.88 | 0.76 | 0.78 | 0.73 | 0.95 | 0.65 | 0.75 |
| Indonesia | 0.89 | 0.81 | 0.59 | 0.76 | 0.81 | 0.59 | 0.74 | 0.76 | 0.86 | 0.56 | 0.66 |
| Iran | 0.90 | 0.87 | 0.57 | 0.79 | 0.87 | 0.83 | 0.80 | 0.82 | 0.91 | 0.63 | 0.84 |

Appendix 2.4(b) continued

| Country study | All questions (ambulatory) | All questions (hospital) | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice of care provider | Quality of basic amenities | Social support | Prompt attention (hospital) |
|----------------------|----------------------------|--------------------------|------------------|---------|---------------|----------|-----------------|-------------------------|----------------------------|----------------|-----------------------------|
| Jordan | 0.95 | 0.84 | 0.56 | 0.85 | 0.92 | 0.85 | 0.89 | 0.76 | 0.91 | 0.57 | 0.89 |
| Morocco | 0.97 | 0.94 | 0.82 | 0.94 | 0.94 | 0.90 | 0.92 | 0.86 | 0.97 | 0.83 | 0.91 |
| Nigeria | 0.90 | 0.91 | 0.69 | 0.87 | 0.86 | 0.86 | 0.83 | 0.83 | 0.94 | 0.79 | 0.78 |
| Oman | 0.90 | 0.81 | 0.40 | 0.77 | 0.83 | 0.82 | 0.76 | 0.62 | 0.86 | 0.36 | 0.76 |
| Romania | 0.91 | 0.85 | 0.62 | 0.80 | 0.85 | 0.72 | 0.81 | 0.79 | 0.91 | 0.61 | 0.76 |
| Russian Federation | 0.92 | 0.85 | 0.64 | 0.87 | 0.86 | 0.71 | 0.81 | 0.81 | 0.84 | 0.69 | 0.75 |
| Syrian Arab Republic | 0.95 | 0.93 | 0.78 | 0.90 | 0.93 | 0.86 | 0.90 | 0.84 | 0.96 | 0.75 | 0.88 |
| Turkey | 0.96 | 0.92 | 0.74 | 0.93 | 0.93 | 0.91 | 0.90 | 0.86 | 0.94 | 0.79 | 0.84 |
| Venezuela | 0.91 | 0.90 | 0.54 | 0.76 | 0.82 | 0.73 | 0.74 | 0.74 | 0.90 | 0.96 | 0.77 |
| Average | 0.93 | 0.88 | 0.65 | 0.84 | 0.88 | 0.81 | 0.83 | 0.81 | 0.92 | 0.69 | 0.81 |
| Standard deviation | 0.03 | 0.04 | 0.11 | 0.06 | 0.04 | 0.08 | 0.06 | 0.07 | 0.04 | 0.14 | 0.07 |

Appendix 2.5 Temporal reliability measured by Kappa statistics

| Question | China | Colombia ^a | Egypt | Georgia | India | Iran | Nigeria ^a | Slovakia | Syria | Turkey ^a |
|----------|-------|-----------------------|-------|---------|-------|------|----------------------|-----------------|-------|---------------------|
| | 409 | 390 | 264 | 246 | 283 | 686 | 58 | 72 | 314 | 132 |
| A1 | 0.79 | 0.34 | 0.82 | 0.41 | 0.62 | 0.43 | 0.47 | 0.76 | 0.67 | 0.87 |
| A2 | 0.68 | 0.38 | 0.83 | 0.34 | 0.53 | 0.46 | 0.29 | 0.77 | 0.72 | 0.83 |
| A3 | 0.78 | 0.47 | 0.84 | 0.36 | 0.59 | 0.49 | 0.39 | 0.83 | 0.72 | 0.83 |
| A4 | 0.76 | 0.32 | 0.86 | 0.32 | | 0.49 | 0.53 | 0.75 | 0.71 | 0.70 |
| A5 | 0.70 | 0.29 | 0.82 | 0.36 | 0.63 | 0.49 | 0.48 | 0.83 | 0.65 | 0.63 |
| A6 | 0.72 | 0.37 | 0.85 | 0.29 | 0.64 | 0.44 | 0.39 | 0.87 | 0.66 | 0.77 |
| A7 | 0.76 | 0.30 | 0.80 | 0.41 | 0.57 | 0.51 | -0.04 | 0.77 | 0.62 | 0.81 |
| A8 | 0.72 | 0.36 | 0.81 | 0.40 | 0.66 | 0.48 | -0.07 | 0.91 | 0.66 | 0.84 |
| A9 | 0.77 | 0.30 | 0.83 | 0.34 | 0.54 | 0.56 | -0.09 | 0.81 | 0.61 | 0.74 |
| A10 | 0.71 | 0.39 | 0.81 | 0.49 | 0.59 | 0.53 | 0.10 | 0.80 | 0.67 | 0.79 |
| A11 | 0.81 | 0.37 | 0.87 | 0.55 | 0.62 | 0.53 | 0.35 | 0.87 | 0.73 | 0.89 |
| A12 | 0.78 | 0.35 | 0.84 | 0.46 | 0.66 | 0.55 | 0.35 | 0.92 | 0.81 | 0.86 |
| A13 | 0.79 | 0.37 | 0.83 | 0.37 | 0.60 | 0.49 | 0.35 | 0.94 | 0.70 | 0.87 |
| A14 | 0.80 | 0.26 | 0.79 | 0.38 | 0.75 | 0.53 | -0.06 | 0.94 | 0.67 | 0.81 |
| A15 | 0.80 | 0.24 | 0.77 | 0.42 | 0.64 | 0.50 | -0.05 | 0.90 | 0.68 | 0.84 |
| A16 | 0.75 | 0.35 | 0.81 | 0.43 | 0.69 | 0.48 | 0.24 | 0.86 | 0.64 | 0.83 |
| A17 | 0.84 | 0.31 | 0.81 | 0.35 | 0.75 | 0.52 | 0.48 | 0.89 | 0.72 | 0.88 |
| A18 | 0.84 | 0.29 | 0.79 | 0.42 | 0.64 | 0.55 | 0.40 | 0.91 | 0.69 | 0.97 |
| A19 | 0.78 | 0.28 | 0.83 | 0.39 | 0.65 | 0.47 | 0.22 | 0.92 | 0.61 | 0.82 |
| A20 | 0.80 | 0.39 | 0.92 | 0.43 | 0.63 | 0.52 | 0.40 | 0.87 | 0.69 | 0.84 |
| A21 | 0.78 | 0.38 | 0.91 | 0.45 | 0.64 | 0.55 | 0.34 | 0.89 | 0.67 | 0.83 |
| A22 | 0.80 | 0.36 | 0.89 | 0.46 | 0.63 | 0.53 | 0.44 | 0.86 | 0.67 | 0.83 |
| Average | 0.77 | 0.34 | 0.83 | 0.40 | 0.63 | 0.50 | 0.27 ^b | 0.86 | 0.68 | 0.82 ^b |
| N | 73 | 53 | 30 | 41 | 36 | 70 | 13 ^b | 15 ^b | 82 | 4 ^b |
| I1 | 0.85 | 0.71 | 0.65 | 1.00 | 0.72 | | | | | |
| I2 | 0.70 | 0.60 | 1.00 | 0.42 | 0.00 | 0.53 | n/a | n/a | 0.84 | n/a |
| I3 | 0.78 | 0.61 | 0.96 | 0.54 | 0.61 | 0.61 | n/a | n/a | 0.86 | n/a |
| I4 | 0.77 | 0.52 | 0.96 | 0.00 | 0.00 | 0.00 | n/a | n/a | 0.00 | n/a |
| I5 | 0.79 | 0.34 | 0.76 | 0.00 | 0.00 | 0.00 | n/a | n/a | 0.00 | n/a |
| I6 | 0.79 | 0.27 | 0.86 | 0.00 | 0.00 | 0.00 | n/a | n/a | 0.00 | n/a |
| I7 | 0.76 | 0.45 | 0.87 | 0.56 | 0.69 | 0.42 | n/a | n/a | 0.68 | n/a |
| I8 | 0.80 | 0.26 | 0.97 | 0.56 | 0.66 | 0.48 | n/a | n/a | 0.77 | n/a |
| I9 | 0.86 | 0.60 | 0.91 | 0.68 | 0.75 | 0.59 | n/a | n/a | 0.85 | n/a |
| I10 | 0.84 | 0.41 | 0.79 | 0.51 | 0.70 | 0.36 | n/a | n/a | 0.72 | n/a |
| I11 | 0.91 | 0.29 | 1.00 | 0.41 | 0.69 | 0.40 | n/a | n/a | 0.85 | n/a |
| I12 | 0.78 | 0.42 | 0.90 | 0.30 | 0.60 | 0.35 | n/a | n/a | 0.77 | n/a |
| Average | 0.80 | 0.46 | 0.89 | 0.42 | 0.45 | 0.34 | | | 0.58 | |

^a Colombia and Turkey did not complete the urban-rural variable. Nigeria had only urban respondents; ^b Excludes Nigeria, Slovakia and Turkey; retest numbers less than 20; n/a means not available

CHAPTER 3

Health systems responsiveness – a measure of the acceptability of health-care processes and systems from the user's perspective



N.B. Valentine, A. Prasad, N. Rice, S. Robone, S. Chatterji
In: P.C. Smith, E. Mossialos, I. Papanicolas, S. Leatherman, editors.
Performance measurement for health system improvement. Cambridge:
Cambridge University Press; 2009. p. 138-86

INTRODUCTION

WHO developed and proposed the concept of responsiveness, defining it as aspects of the way individuals are treated and the environment in which they are treated during health system interactions.¹ The concept covers a set of non-clinical and non-financial dimensions of quality of care that reflect respect for human dignity and interpersonal aspects of the care process, which Donabedian (1980) describes as “the vehicle by which technical care is implemented and on which its success depends”.² Eight dimensions (or domains) are collectively described as goals for health care processes and systems (along with the goals of higher average health and lower health inequalities; and non-impoverishment—as measured through other indicators): (i) dignity, (ii) autonomy, (iii) confidentiality, (iv) communication, (v) prompt attention, (vi) quality (of) basic amenities, (vii) access to social support networks during treatment (social support), and (viii) choice (of health care providers).

Building on extensive previous work, this chapter directs the conceptual and methodological aspects of the responsiveness work in three new directions. First, the given and defined domains are used to link responsiveness (conceptually and empirically) to the increasingly important health systems concepts of access to care and equity in access. The concept of equity used in this chapter was defined by a WHO working group with experts on human rights, ethics and equity. It is defined as the absence of avoidable or remediable differences among populations or groups defined socially, economically, demographically, or geographically. Health inequities involve more than inequality – whether in health determinants or outcomes, or in access to the resources needed to improve and maintain health. They also represent a failure to avoid or overcome such inequality which infringes human rights norms or is otherwise unfair.³ Second, it expands on the issue of measurement strategies. Third, the psychometric results of the responsiveness module from the World Health Survey (WHS) are compared with its survey instrument predecessor in the Multi-Country Survey (MCS) Study.

The chapter concludes with analysis of the most recent results for responsiveness from the WHS for ambulatory and inpatient health care services for 65 countries (with special reference to subsets of European countries) to see how European countries' health care systems perform with respect to responsiveness.

RESPONSIVENESS OPERATIONALIZED AS A POPULATION HEALTH CONCEPT

Responsiveness is measured using criteria related to the importance of users' views. Individuals who use (or decide not to use) the health care system are viewed as the appropriate source of information on non-technical aspects of care. This approach implies measuring responsiveness through household or other types of user surveys rather than, for example, expert opinion or facility audits.

Concepts such as quality of life and general satisfaction are also measured in surveys. However, self-reports have the additional criterion that they should be linked to one or several actual experiences with health services in the respondent's recent past (previous year) and upon which they base their views. These experiences are usually based on some type of interaction with the health care system including interaction with a specific person in that system; a communication campaign; or another type of health system event or action that did not entail direct personal interactions. This criterion places the focus on what actually happened during contact with the health care system, rather than the respondent's satisfaction or expectations regarding the health care system in general.

WHO (2000)⁴ broadly defines the health care system as all actions whose primary intent is to produce health. The responsiveness measure proposed by WHO conceptually aims to measure the responsiveness of the 'whole health care system' to the 'whole population'.⁵ When the self-report measurement approach based on the criterion of an actual (recent) experience is combined with the concept of measuring the whole population's experience of the whole health care system then the measurement challenges are multiplied. We outline aspects of these challenges below.

Spheres of health events

Seven different types of health events that require interactions with health care systems or services are listed below. The list is intended to be relevant generically, regardless of the configuration of providers, financing, technology, medicines and human resources:

1. ambulatory care in response to acute needs;
2. ambulatory care for chronic conditions;
3. inpatient care for short-term stays (>24 hours; <3 months);
4. long-term institutionalized care: e.g., for populations with mental illnesses, disabilities related to physical health conditions or elderly populations;

5. non-excludable public health interventions: e.g., public health promotion for communities or population groups such as access to improved water and sanitation, smoking bans;
6. opportunities for participation in health system governance: e.g., shaping the health system and issues affecting health;
7. administrative and financial transactions: e.g., ease of making payments for services and medicines or of obtaining medicines with prescriptions, receiving reimbursement from insurance if needed.

The list illustrates that the design of questions in household or user surveys and the actual survey coverage would require significant work to cover the entire typology of interactions and abide by the criterion of obtaining user reports. For example, individuals receiving long-term institutionalized care cannot respond to household surveys and require more targeted designs. Also, questions may need to be tailored to the specific institutional arrangements of services (including insurance coverage) for the country, region or sector.

Role of the users

Given that the health care system is a socially constructed system, individuals' interactions with that system will differ according to circumstances. These can be categorized into four non-mutually exclusive groupings. For any given time period, a single survey respondent may have experiences of interactions that relate to all, none or some of these roles:

- a. a patient or user (with or without personal contact);
- b. a patient or user by proxy e.g., chiefly for children, but also for people with mental illness or elderly persons;
- c. a relative or a close friend of a patient;
- d. a member of society who uses health services, but has not done so in the defined period of the previous year, and who has some ability to shape the structure of health institutions. This citizen role is facilitated by the mechanisms for social participation in decision-making on health.

Combining health events and user roles - interactions

The full range of interactions combines user roles and different types of health events. When these are stated explicitly they help policy-makers to understand which aspects of responsiveness they are most interested in capturing. A strategy to measure all these combinations of interactions and user roles would need to identify the most important in order to avoid overburdening respondents. This breadth of responsiveness is operationally challenging and to date has

not been undertaken systematically in any country. Nevertheless, from a heuristic point of view, it is important to observe the potential implications of a concept if operationalized fully. It is also vital to decide whether measurement is necessary for all domains of responsiveness or only one. WHO designed the WHS responsiveness instrument to cover interactions represented by the combination of events and user roles matching the alphanumeric labels listed above – 1ab, 2ab, 3ab, and 6d (involvement in decision making only).

Responsiveness and equity in access

The link between responsiveness and equity in access is important. It derives from the impact of service qualities described by the responsiveness domains on utilization patterns. An explicit framework that describes how responsiveness is linked to access to care via the care context and process can inform empirical work aimed at describing responsiveness across countries. Figure 3.1 presents such a framework that builds on other frameworks in the literature covering the medical-care process⁶; access to care^{7,8}; utilization (Andersen's 1995 model⁹ as described by Bradley et al.¹⁰); and the conceptual framework proposed to the Commission on Social Determinants of Health.¹¹

The framework has three broad components: (i) environment; (ii) agents defining need for care; and (iii) process of care and outcomes. The first two components delineate context and together define the need for care at the population level. Their development was informed by the Aday and Andersen framework⁷ of 'health policy'; 'population characteristics'; 'health service characteristics'; and 'utilization', with some adaptations. For example, the decision-making agents component in the Figure 3.1 framework draws attention to the role of both providers and users in defining need and setting the context for utilization. It evokes three agency groupings: (i) the providers and their accepted protocols (which may differ across countries); (ii) lay persons (with their socially accepted protocols/norms); and (iii) the specific epidemiological or biological agents which produce different responses from the other two groups of decision-makers.

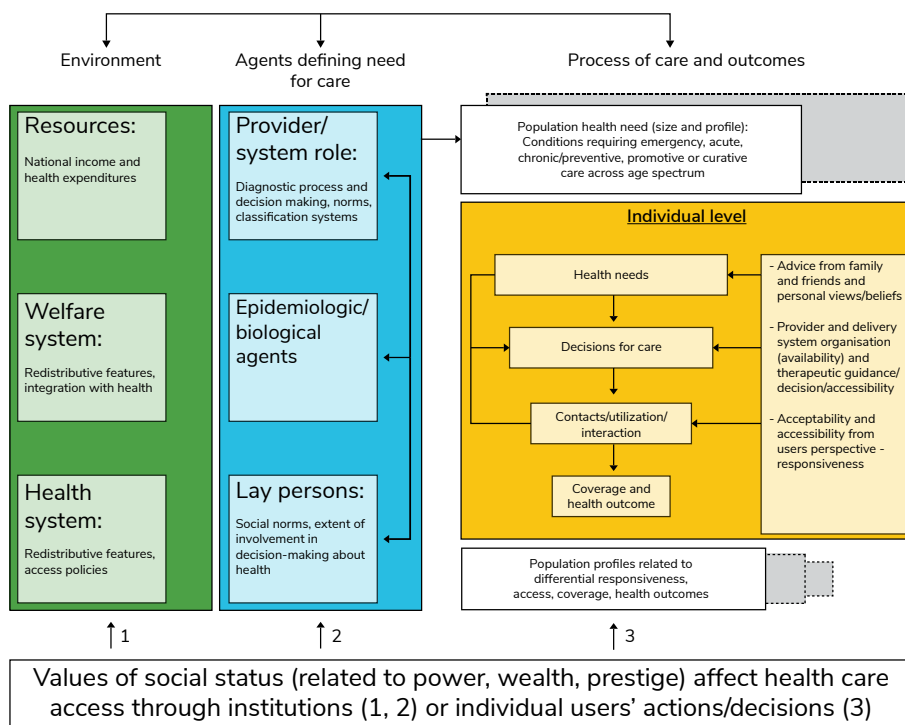


Figure 3.1 Framework for understanding the link between health system responsiveness and equity in access

Recognition of the separate groupings of providers and lay persons is an important innovation that was raised in the Solar and Irwin (2007/2010)¹¹ framework and the work of the Health Systems Knowledge Network of the Commission on Social Determinants of Health.¹² This distinction is important for understanding the context in which responsiveness is measured and the implications for policy discussions. Responsiveness reports on convenience of access or confidentiality will reflect different profiles of services which have been negotiated by decision-making groups in society. For example, midwives in one country may make home visits that are not part of population health needs in another. Differences are to be expected and may provide explanations for varying responsiveness across countries. However, it is important that these factors are explicit in analytical frameworks in order to understand how to improve responsiveness across different countries.

The third component of the framework is most relevant to the measurement of responsiveness – the process of care and outcomes. An individual who has a specific need for care moves from (a) recognition of health needs, to (b) decisions for care, to (c) contact with the system/utilization,

and to (d) coverage. The latter is defined as the single, multiple, or perpetual contacts to ensure adherence that may be required to guarantee adequate care for a particular condition (adapted from Tanahashi⁸). Care-seekers' decisions related to utilization and the possible achievement of full coverage (explained below) for a particular condition are influenced by three broad factors shown in Figure 3.1: (i) the personal context (advice from family and friends, personal beliefs), (ii) providers (administering therapeutic guidelines/decisions, organization of delivery e.g., being able to see a general doctor or specialist directly); and (iii) by the health system's capacity to be responsive. The responsiveness domains mostly relate to Tanahashi's⁸ definitions of accessible (users able to reach and use the health service) and acceptable care or coverage (users willing to use accessible services).

The concept of full coverage is introduced into the framework as coverage, although this term is used infrequently in the traditional 'access' literature (except Tanahashi⁸ and, more recently, Shengelia et al.¹³). It usefully communicates the concept of a norm related to interventions for particular conditions. This differs from utilization rates for which high or low values indicate only the use of health care resources without explicit reference to norms or need related to particular conditions. Health outcomes are affected by the extent of coverage reached and may not be affected by utilization rates. Of course, there is room for both concepts in the same framework as utilization rates for which the vulnerability of the population group is proxied (e.g., by income) do give some indication of the resources consumed relative to need.

The literature does make reference to definitions of coverage at population and individual levels. Shengelia et al.¹³ define effective coverage at the individual level as "the fraction of maximum possible health gain an individual with a health care need can expect to receive from the health system". Tanahashi⁸ refers to a population level measure of coverage as "the number of people for whom the service has satisfied certain criteria relating to its intended health intervention, compared with the total target population".

The third component of the framework also shows the links between responsiveness and equity in access. Responsiveness affects access at the individual level first. Responsiveness that is systematically worse for certain social groups with the same or greater need than other social groups could lead to inequities in access. These are defined as arising when anticipated, perceived or actual responsiveness attributes of the service dissuade certain social groups from seeking care and receiving adequate careⁱ. By adapting Tanahashi's⁸ population-level definition of coverage to the individual level, 'adequate care' would refer to services striving to meet a pre-defined technical norm in response to a variety of health conditions (completion of treatment; or continued, on-going treatment for chronic or palliative cases). Given this relationship between responsiveness, equity in responsiveness and equity in access, it is possible to use measures of

responsiveness inequalities by different social groups (stratified according to need e.g., proxied by income) to anticipate inequities in access.

Equity considerations for responsiveness survey design

A service that is perceived to have poor responsiveness may not be used optimally (or even at all) or as required by the health condition. Yet responsiveness measurement needs to be based on actual interactions. Thus, one weakness of the measurement approach is that measures will be biased upwards, not only because self-reports of this nature are usually biased upwards¹⁴ but also because they do not fully capture the experiences of respondents who are in need but have not used services recently. Responsiveness measurement will not record the experience of care of someone who is excluded from care by failing to initiate⁷ or obtain contact with the system.⁸

Figure 3.2 illustrates how populations may be excluded, with reference to two types of problems. In some cases, populations may not have sought care in the defined time period due to responsiveness or other factors e.g., financial barriers. These denied users would be excluded by screening questions on when they last came into contact with a health service. In other cases, the very nature of their vulnerability (e.g., homelessness) may put certain populations beyond the reach of traditional survey techniques. In both instances, surveys will bias responsiveness upwards and potentially underestimate inequalities in responsiveness. For the first problem, denied users can be asked about the barriers to care in order to gain qualitative information on the responsiveness measures. The second problem will require special efforts (e.g., surveys of institutionalized, homeless or migratory populations).

Special consideration should be given to the inclusion of service contacts with children as exposures at early stages in the life course have not only equity impacts that transmit into adulthood, but also intergenerational consequences. Minors cannot report for themselves but reporting by parents has been shown to be effective. This was used for children up to the age of twelve in the WHS, as recommended by experts.¹⁶

Some critics have argued for special attention on sicker populations¹⁷ to ensure equity and because they are the ones who know the services better. A strategy focusing on the sick may use health facility, exit-based surveys rather than household surveys, although this approach may omit those sick who have not used health services.

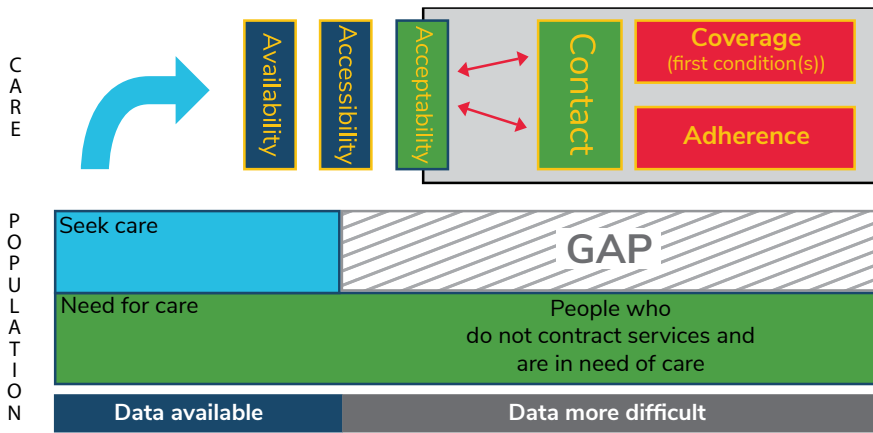


Figure 3.2 Traditional survey methods omit data from certain population groups, overestimating responsiveness

Source: World Health Organization 2008/2010¹⁵

RESPONSIVENESS QUESTIONNAIRES

The responsiveness domains were derived from existing patient questionnaires and studies as reported in the extensive literature review conducted by de Silva.¹⁸ This review profiled the questionnaire work undertaken by the Agency for Healthcare Research and Quality (AHRQ), Harvard Medical School, the Research Triangle Institute and the RAND Corporation. None of the existing questionnaires and studies captured all the dimensions that they covered collectively. WHO developed an instrument (questionnaire) that covered the collection of dimensions (described in the literature review) related to non-technical aspects of the process of care: dignity, autonomy, communication, confidentiality, prompt attention (related to convenience and peace of mind rather than urgent medical attention), quality of basic amenities, access to social support networks during treatment (labelled “social support” in the MCS Study and “access to family and community support” in the World Health Survey), and choice (of health care providers).

WHO’s responsiveness questionnaire has been developed and refined. Questions (items) were initially fielded in a key informants’ survey of 35 countries and the results described in the World Health Report 2000.⁴ A household survey instrument which included pre-testing was then developed as part of the MCS Study covering 60 countries.^{19,20} Following the launch of the MCS Study, the concept of responsiveness and the questionnaire were refined and a revised instrument was included in the World Health Survey (WHS) which was administered across 70

countries in 2002-2003. The WHS basic survey mode used an in-person interview conducted in one of three possible forms: ninety-minute in-household interview (53 countries) (long form), thirty-minute face-to-face interview (short form) (13 countries); or computer-assisted telephone interview. Samples were randomly selected (those above 18 years) resulting in sizes of between 600 and 10,000 for each country surveyed. Descriptive statistics about individuals sampled in each country are reported in Appendix 3.2.

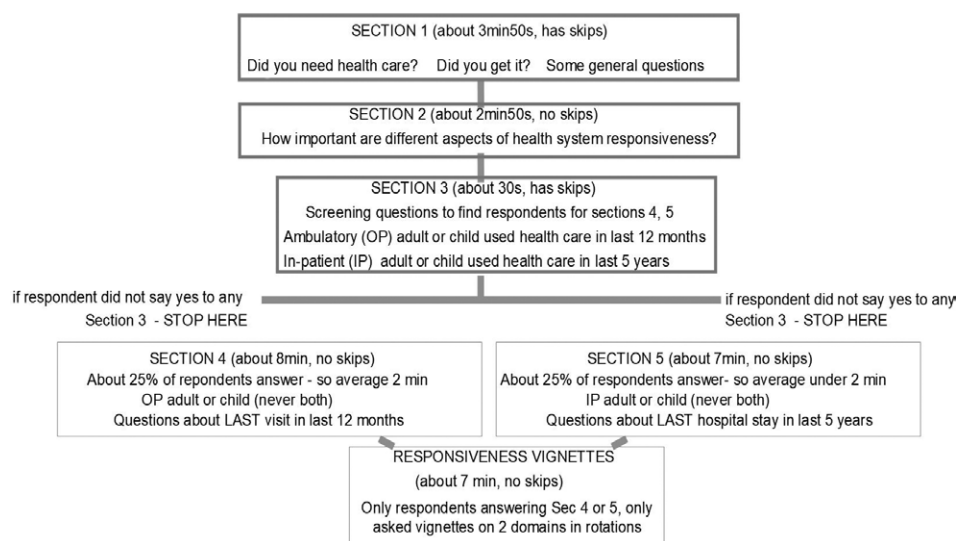


Figure 3.3 Responsiveness questionnaire as a module in the World Health Survey questionnaire: interview structure and timing

Data collection was performed on a modular basis, addressing different aspects of health and the health system and including information on health insurance, health expenditures, socio-demographics and income, health state valuations, health system responsiveness, and health system goals.²¹ Figure 3.3 provides an overview of the responsiveness module in the WHS. The measurement of responsiveness was obtained by asking respondents to rate their most recent experience of contact with the health system within each of the 8 domains by responding to the set of questions listed in Figure 3.4. The response categories available were: very good, good, moderate, bad and very bad.

Like health, responsiveness is viewed as a multidimensional concept. Each domain is measured as a categorical variable for which there is an assumed underlying latent scale. Certain domains are more suited to patient evaluation e.g., quality of basic amenities and prompt attention. It is

more difficult to evaluate whether full details of the nature of the illness and all relevant treatments and available options have been disclosed as this requires specialist knowledge. Accordingly, it is more problematic to maintain objectivity in the evaluation of some domains. Samples have undergone extensive quality assurance procedures both at data collection stage at country and inter-country levels.

The MCS Study and WHS modules on responsiveness have strong similarities. However, they have a number of different ways of expanding and alleviating the burden on survey respondents. More notable changes in the WHS include: more face-to-face interviews or computer assisted telephone interviews (the MCS Study included 28 postal surveys); eliciting the experiences of children up to twelve (reported through a parent); and reducing the number of items that individuals are required to respond to on each domain. The WHS module also tried to identify barriers to access by asking people if they needed care and, if so, whether they sought care or why they did not (Figure 3.3 section 1). The analyses that follow focus on the questions asked in sections four and five of the responsiveness module and cover ambulatory and hospital (inpatient) experiences of adult and children populations.

| Responsiveness domain label (Short description) | Item questions |
|--|---|
| Prompt attention (Convenient travel and short waiting times) | How would you rate: 1- the traveling time to the hospital 2- the time you waited before being attended to ^b |
| Dignity (Respectful treatment and communication) | How would you rate: 1- being greeted and talked to respectfully ^a 2- the respect for privacy during physical examinations and treatments ^{a b} |
| Communication (Clarity of communication) | How would you rate: 1- how clearly health care providers explained things to you ^a 2- the time you get to ask questions about your health problems or treatment ^{a b} |
| Autonomy (Involvement in decisions) | How would you rate: 1- being involved in making decisions about your health care or treatment ^a 2- the information you get about other types of treatments or tests ^b |
| Confidentiality (Confidentiality of personal information) | How would you rate the way: 1- the health services ensured you could talk privately to health care providers ^a 2- your personal information was kept confidential ^{a b} |
| Choice (Choice of health care provider) | How would you rate: 1- the freedom you had to choose the health care providers that attended to you |
| Quality of basic amenities (Surroundings/facilities) | How would you rate: 1- the cleanliness of the rooms inside the facility, including toilets ^a 2- the amount of space you had ^{a b} |
| Access to family and community support (Contact with outside world and maintenance of regular activities/ social support) | How would you rate: 1- the ease of having family and friends visit you 2- the experience of staying in contact with the outside world when you were in hospital ^{a b} |

Figure 3.4 Operationalization of the responsiveness domains in the WHS

^a Similar items appear in the Multi-Country Survey Study; ^b Item dropped for the short version of the World Health Survey

PSYCHOMETRIC PROPERTIES OF THE RESPONSIVENESS DOMAIN QUESTIONS

Psychometrics examines the quality of survey instruments, and has been used extensively to assess the quality of the responsiveness instrument in both the MCS Study and the WHS. This section briefly considers three key desirable properties of a survey instrument (feasibility, reliability and validity) and compares them in the MCS Study and the WHS. The results on these properties are presented in combination for ambulatory and home care (as ambulatory care) and separately for inpatients. A more detailed description of the psychometric properties of the MCS Study is provided by Valentine et al.^{20,22}

Feasibility

Feasibility refers to the ease of administering an instrument in the field and can be assessed by considering factors such as survey response rates, the proportion of missing items in a respondent interview (inappropriate missing responses) and item missing rates (percentage of respondents who omitted a particular item). The literature provides little indication of an acceptable survey response or inappropriate response missing rates but, in general, guidance indicates that item missing rates below 20% can be considered acceptable.^{20,23}

Survey response rates measured as a percentage of attempted and effective contacts were available only for the MCS. The comparison of reliability between the two surveys rests mainly on interview completion (a form of survey response rates) and item missing rates. It is important to note that completion rates may be as high as 100% as they give the number of persons who started and completed interviews as a percentage of the number of persons who started interviews.

The MCS Study shows high measures of feasibility with a response completion rate greater than 95% for each of the countries considered, exception Colombia (73%). Furthermore, no country exceeded a 20% item missing rate and only 3 countries had item missing rates in excess of 10% (Switzerland, Turkey and Tobago). Valentine et al.²⁰ provide full results of the psychometric properties of the MCS Study. A similar analysis of the responsiveness instrument in the WHS showed that response completion rates per country were greater than 80% for all countries with the exception of Israel (63%). No country exceeded the accepted item missing rate threshold of 20% for ambulatory care, while only Swaziland exceeded this threshold for inpatient care.

Additional information on the feasibility of the WHS responsiveness instrument is provided by the percentage of respondents that report missing values for zero; one; two; or three or more items. In countries where the long-form questionnaire was implemented, in responses on am-

bulatory care 88% of respondents returned no missing items; 6% reported one; 2% reported two and 4%, three or more. Corresponding values for inpatient care were 87%, 5%, 1% and 7%. In countries where the short form questionnaire was implemented, in responses for ambulatory care 87% returned no missing items, 11% reported one; 3% reported two and 2% reported three or more. The corresponding figures for inpatient care are 81%, 11%, 4% and 4%.

Table 3.1 Item missing rates, ambulatory care (%)

| | MCS Study | WHS |
|------------------|-----------|------|
| Prompt attention | 0.86 | 1.72 |
| Dignity | 1.13 | 1.75 |
| Communication | 0.55 | 0.38 |
| Autonomy | 2.70 | 2.03 |
| Confidentiality | 6.40 | 2.43 |
| Choice | 7.50 | 3.25 |
| Basic amenities | 2.30 | 3.25 |
| Average | 3.06 | 2.12 |

Table 3.1 offers a more direct comparison of the item missing rates. The values for the MCS Study are taken from Valentine et al.²⁰ and consider only the 41 countries where interviewer administered interviews were held, corresponding to the method used in the WHS. Item missing rates are provided for ambulatory care by domain (calculated as the arithmetic mean of missing rates of individual items present in a domain) by averaging across countries. As can be seen, the WHS reported lower item missing rates for 4 of the 7 domains and failed to exceed 3.25% in any domain. Averaged across countries and domains, the overall missing item rate is nearly 1% lower than in the MCS Study.

Reliability

The reliability of an instrument refers to the test-retest property of measurement, usually over time, all other things being equal. Temporal reliability can be measured using the Kappa statistic. Landis and Koch²⁴ suggest that statistics in the range 0.41–0.60 indicate moderate reproducibility; 0.61–0.80 substantial reproducibility and 0.81–1.00 almost perfect reproducibility. Table 3.2 compares Kappa statistic results for the MCS Study and WHS.

Instrument reliability in the MCS Study was assessed by re-administering the entire responsiveness questionnaire to respondents in 10 country sites one month after the initial interview. There

is a high reliability of all items by domain when averaged across the countries.²⁰ The lowest Kappa value reported for any domain was 0.64 (for dignity in home care). However, there is variability in reliability when results are averaged across domains within countries. Reproducibility is substantial in 5 countries, moderate in 3 and low in 2.

The reliability of the WHS instrument was assessed by re-interviewing 10% of the original sample in each country. The re-interviewed respondents were selected randomly and asked to complete the follow-up questionnaire one to seven days following the first interview.²⁵ We consider reliability in the 53 countries for ambulatory care, and the 55 countries for inpatient care where sufficient data points (>20) were available in the follow-up survey. When the Kappa statistics are averaged across items within countries, at least moderate reliability was reported for ambulatory care in 24 countries and for inpatient care in 27 countries. When results are averaged across countries for each item separately all items satisfy at least the condition for moderate reproducibility.

Table 3.2 compares Kappa statistics for the MCS Study and the WHS. The Kappa statistic is provided for each domain, averaged across countries and overall for countries and domains. The first and second columns in Table 3.2 show Kappa statistics averaged across the 10 countries in the MCS Study and the 53 countries of the WHS in which the responsiveness instrument was re-administered to respondents. When considering all available countries, the Kappa statistics are considerably lower for the WHS. However, this does not provide a like-for-like comparison. Consideration of the 2 countries common to both surveys (India and China) provided in columns 3 and 4 indicates very similar reliability in each survey.

Table 3.2 Reliability in the MCS Study and WHS

| | MCSS ^a (10 countries) | WHS (53 countries) | MCSS ^a (India, China) | WHS (India, China) |
|------------------|----------------------------------|--------------------|----------------------------------|--------------------|
| Prompt attention | 0.60 | 0.49 | 0.66 | 0.73 |
| Dignity | 0.61 | 0.45 | 0.69 | 0.71 |
| Communication | 0.57 | 0.45 | 0.67 | 0.73 |
| Autonomy | 0.65 | 0.46 | 0.71 | 0.70 |
| Confidentiality | 0.59 | 0.45 | 0.74 | 0.71 |
| Choice | 0.63 | 0.40 | 0.75 | 0.72 |
| Basic amenities | 0.65 | 0.44 | 0.71 | 0.72 |

^a Source: Valentine et al. 2007²⁰

Psychometric measures can be investigated where data are stratified by population groups of interest. This allows an assessment of whether any revealed systematic variations suggest caution in interpreting results or indicate a need for greater testing before a survey is implemented.

We investigated the reliability of the WHS responsiveness instrument across European countries for two population groups defined by educational tenure. Table 3.3 presents average Kappa statistics for each domain separately for western European countries and countries of Central and Eastern Europe and the former Soviet Union (CEE/FSU) (listed in Appendix 3.1). Results are further presented by level of educational tenure (defined as people having studied for more or less than twelve years). Table 3.3a and Table 3.3b report results for ambulatory and inpatient care, respectively. Overall, the reliability of the responsiveness instrument appears to be greater in CEE/FSU countries than in western European countries, irrespective of levels of education.

Interestingly, country groupings indicate that the reliability of the instrument is greater for less educated individuals in CEE/FSU countries but generally the opposite appears to hold for Western Europe. Taken in their totality across both groups of countries, the results suggest that (with the exception of the domain for confidentiality and choice) educational achievement has little influence on the reliability of the responsiveness instrument. Further, the reliability of the instrument for ambulatory care appears marginally better than for inpatient care (except for quality of basic amenities domain).

Table 3.3 Reliability across European countries

a. Ambulatory care

| | Western Europe | | CEE/FSU | | Europe overall | |
|------------------|----------------|----------------|---------------|----------------|----------------|----------------|
| | Low Education | High Education | Low Education | High Education | Low Education | High Education |
| Prompt attention | 0.49 | 0.44 | 0.59 | 0.56 | 0.54 | 0.50 |
| Dignity | 0.40 | 0.40 | 0.57 | 0.60 | 0.49 | 0.50 |
| Communication | 0.42 | 0.42 | 0.52 | 0.49 | 0.47 | 0.45 |
| Autonomy | 0.43 | 0.41 | 0.55 | 0.46 | 0.49 | 0.43 |
| Confidentiality | 0.25 | 0.52 | 0.58 | 0.52 | 0.41 | 0.52 |
| Choice | 0.37 | 0.26 | 0.61 | 0.52 | 0.49 | 0.39 |
| Basic amenities | 0.24 | 0.37 | 0.54 | 0.53 | 0.39 | 0.45 |
| Average | 0.37 | 0.40 | 0.56 | 0.52 | 0.47 | 0.46 |

Table 3.3. Reliability across European countries (continued)**b. Inpatient care**

| | Western Europe | | CEE/FSU | | Europe overall | |
|------------------|----------------|----------------|---------------|----------------|----------------|----------------|
| | Low Education | High Education | Low Education | High Education | Low Education | High Education |
| Prompt attention | 0.30 | 0.38 | 0.68 | 0.53 | 0.49 | 0.45 |
| Dignity | 0.34 | 0.40 | 0.65 | 0.53 | 0.50 | 0.47 |
| Communication | 0.25 | 0.34 | 0.56 | 0.52 | 0.41 | 0.43 |
| Autonomy | 0.19 | 0.24 | 0.61 | 0.48 | 0.40 | 0.36 |
| Confidentiality | 0.21 | 0.37 | 0.60 | 0.49 | 0.41 | 0.43 |
| Choice | 0.23 | 0.34 | 0.64 | 0.49 | 0.43 | 0.42 |
| Basic amenities | 0.29 | 0.43 | 0.62 | 0.52 | 0.46 | 0.47 |
| Social support | 0.26 | 0.38 | 0.60 | 0.49 | 0.43 | 0.43 |
| Average | 0.26 | 0.36 | 0.62 | 0.51 | 0.44 | 0.43 |

Validity

The psychometric property of validity focuses on exploring the internal structure of the responsiveness concept, particularly the homogeneity or uni-dimensionality of responsiveness domains. The property is often measured through factor analysis and Cronbach's alpha. Stronger evidence of uni-dimensionality (factor loadings close to +1 or -1) supports greater validity of the instrument; a minimum value in the range of 0.6 to 0.7, has been suggested for Cronbach's alpha.^{26,27}

Validity was assessed by pooling data from different countries and analysing each domain independently. For the MCS Study, values of Cronbach's alpha suggested that all domains lay within the desired range and were greater than 0.7 for all except one (prompt attention = 0.61).²⁰ For the WHS all countries satisfied the requirement that Cronbach's alpha is greater than 0.6 – the minimum value across countries was 0.66 for inpatient care and 0.65 for ambulatory care. This requirement was also satisfied for all domains except prompt attention for ambulatory care (alpha = 0.56).

We further evaluated the construct validity of the WHS questionnaire using maximum likelihood exploratory factor analysis, as performed by Valentine et al.²⁰ when analysing the MCS Study ambulatory responsiveness questions (the inpatient sector of MCS Study only contained 1 item per domain, except for prompt attention and social support). The method makes reference to Kaiser's eigenvalue rule which stipulates that item loadings on factors should be 0.40 or greater.²⁸ The results of MCS Study analysis are presented by Valentine et al.²⁰

Valentine et al.'s results confirmed the hypothesized domain taxonomy for the majority of the domains.²⁰ The high human development countries have a few exceptions within the domains of prompt attention and dignity, where items tend to load on multiple factors. For the WHS questionnaire, Table 3.4 reports the promax rotated factor solutions for ambulatory care computed across all countries (pooled) in which the long-form questionnaire was implementedⁱⁱ. In general, results confirmed the hypothesized domain taxonomy, as the items belonging to particular domains loaded onto a single factor, except in the case of autonomy. For autonomy, the factor for communication had the largest loading for the first item but the second largest loading (0.371) corresponded to the largest loading on the second item (factor 5). For prompt attention, the two largest loadings fell onto a single factor (7) but did not reach the threshold suggested by Nunnally and Bernstein.²⁸

Table 3.4 Promax rotated factor solution for the ambulatory responsiveness questions in the WHS

| Domain | Item | Latent underlying factor | | | | | | | Uniqueness |
|--------------------------------|------|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|------------|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Prompt attention | 1 | -0.018 | 0.115 | 0.135 | -0.006 | 0.056 | -0.013 | 0.288 | 0.774 |
| | 2 | 0.010 | -0.019 | -0.038 | 0.013 | -0.019 | 0.019 | 1.023 | 0.000 |
| Dignity | 1 | 0.048 | 0.728 | 0.045 | -0.046 | 0.044 | -0.027 | 0.061 | 0.352 |
| | 2 | 0.025 | 0.719 | -0.079 | 0.225 | -0.009 | -0.003 | -0.041 | 0.311 |
| Communication | 1 | 0.523 | 0.321 | -0.063 | 0.076 | 0.048 | 0.014 | -0.014 | 0.327 |
| | 2 | 0.855 | -0.017 | 0.038 | 0.000 | 0.048 | 0.011 | 0.019 | 0.157 |
| Autonomy | 1 | 0.476 | -0.027 | 0.042 | 0.020 | 0.371 | 0.034 | 0.021 | 0.294 |
| | 2 | -0.011 | 0.029 | 0.010 | -0.005 | 0.924 | 0.028 | -0.017 | 0.116 |
| Confidentiality | 1 | 0.072 | 0.039 | -0.030 | 0.614 | 0.194 | 0.013 | 0.032 | 0.327 |
| | 2 | -0.005 | 0.028 | 0.033 | 0.849 | -0.050 | -0.005 | 0.010 | 0.257 |
| Choice | 1 | 0.072 | -0.055 | 0.629 | 0.037 | 0.145 | -0.042 | -0.038 | 0.462 |
| | 2 | -0.021 | 0.169 | 0.185 | 0.134 | -0.058 | 0.444 | -0.043 | 0.462 |
| Basic amenities/ facilities | 1 | -0.021 | 0.169 | 0.185 | 0.134 | -0.058 | 0.444 | -0.043 | 0.462 |
| | 2 | 0.016 | -0.050 | -0.063 | -0.028 | 0.034 | 1.052 | 0.026 | 0.000 |

As seen in Table 3.5, the hypothesized domain taxonomy was also confirmed for inpatient care and, again, the items failed to load on a single factor in only two domains (prompt attention and communication). The communication item related to information exchange loaded more strongly on the autonomy domain. In general, the strong association between autonomy, communication and dignity domain items supports the assertions made in previous MCS Study work and

Table 3.5 Promax rotated factor solution for the inpatient responsiveness questions in the WHS

| Domain | Item | Latent underlying factor | | | | | | | | | | Uniqueness |
|----------------------------|------|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------|--------------|--------|------------|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| Prompt attention | 1 | 0.009 | 0.002 | -0.073 | -0.011 | 0.005 | -0.004 | -0.011 | -0.011 | 1.041 | 0.007 | 0.000 |
| | 2 | -0.007 | -0.004 | 0.446 | 0.063 | -0.021 | 0.031 | 0.051 | 0.044 | 0.233 | -0.037 | 0.543 |
| Dignity | 1 | 0.036 | -0.051 | 1.007 | -0.023 | -0.018 | -0.007 | 0.014 | -0.012 | -0.081 | 0.005 | 0.134 |
| | 2 | 0.052 | 0.263 | 0.437 | 0.008 | 0.172 | 0.024 | -0.099 | -0.002 | 0.010 | 0.029 | 0.371 |
| Communication | 1 | 0.150 | -0.016 | 0.038 | 0.004 | 0.786 | 0.005 | 0.019 | 0.022 | 0.009 | -0.005 | 0.131 |
| | 2 | 0.526 | -0.002 | 0.032 | 0.003 | 0.144 | 0.015 | 0.025 | 0.292 | -0.012 | -0.001 | 0.239 |
| Autonomy | 1 | 0.757 | 0.040 | -0.009 | 0.028 | -0.021 | 0.009 | -0.030 | 0.167 | -0.002 | 0.002 | 0.253 |
| | 2 | 0.951 | -0.011 | 0.046 | -0.004 | 0.009 | 0.010 | 0.017 | -0.219 | 0.028 | -0.004 | 0.184 |
| Confidentiality | 1 | 0.178 | 0.632 | 0.011 | 0.098 | 0.032 | 0.010 | 0.028 | -0.022 | -0.034 | -0.134 | 0.307 |
| | 2 | 0.026 | 0.874 | -0.016 | -0.055 | -0.033 | 0.017 | 0.009 | 0.014 | 0.021 | 0.013 | 0.269 |
| Choice | 1 | 0.254 | 0.053 | 0.006 | 0.007 | 0.021 | 0.024 | 0.475 | 0.007 | -0.017 | 0.012 | 0.455 |
| Basic amenities/facilities | 1 | 0.026 | 0.091 | 0.060 | 0.501 | 0.004 | 0.034 | 0.067 | -0.013 | 0.019 | 0.141 | 0.417 |
| | 2 | 0.017 | -0.045 | -0.037 | 0.959 | -0.002 | 0.035 | -0.032 | 0.007 | -0.014 | 0.007 | 0.147 |
| Social support | 1 | -0.014 | 0.031 | 0.029 | 0.121 | 0.016 | 0.747 | -0.027 | -0.019 | -0.011 | -0.003 | 0.294 |
| | 2 | 0.039 | -0.011 | -0.021 | -0.034 | -0.010 | 0.871 | 0.024 | 0.016 | 0.006 | 0.003 | 0.244 |

elsewhere that communication is an important precondition or accompaniment to being treated with dignity and involvement in decision-making about care or treatment.

MEASURING RESPONSIVENESS

Calculating the measures

Two measures are used to capture health system responsiveness in the analyses that follow. The first is the level of responsiveness; the second is the extent of inequalities in responsiveness across socio-economic groups in a country. This second measure can be used as a proxy for equity in responsiveness as will be explained below. Both measures are applied to user reports from ambulatory and inpatient health care settings, resulting in four indicators per country.

The level of responsiveness (also called the responsiveness score) is calculated by averaging the percentage of respondents reporting that their last interaction with the health care system was good or very good across the relevant domains (7 domains for ambulatory care; 8 domains

for inpatient). This average is referred to as overall ambulatory or inpatient responsiveness. A higher value indicates better responsiveness. Scores or rates per country are age-standardized using the WHO World Standard Population table, given that increasing age is associated with increasingly positive reports of experiences with health services.²⁹

The inequality measure is based on the difference across socio-economic groups, in this case identified by income quintiles and a reference groupⁱⁱⁱ.³⁰ From a theoretical perspective, the reference group could be chosen on the basis of the best rate in the population; the rate in the highest socioeconomic group; a target external rate; or the mean rate of the population. The highest income quintile reference group was selected here. Each difference between the highest and other quintiles is weighted by the size of the group with respect to the reference group. The measure is calculated for each domain and then an average is taken across all domains to derive a country inequality indicator (again, for ambulatory or inpatient services separately)^{iv}. Higher value for the inequality measure indicates higher inequalities and, by proxy, higher inequities (see below).

The assumption behind the link between the inequality measure of responsiveness calculated here and an inequity measure is based on the equity criterion that there should be an equal level of responsiveness for people with equal levels of health need. To the extent to which income may proxy health needs (assuming a negative relationship between income and ill-health), then a positive gradient between income quintiles and responsiveness levels provides evidence of inequity. In other words, a positive gradient in responsiveness from low to high income groups would imply inequities in responsiveness. Lower income groups would presumably have greater health service needs and be entitled to at least the same, or better, responsiveness from the health system.

All domain results were sample weighted and average responsiveness scores were age-standardized because of the widespread evidence in the literature of a systematic upward bias in ratings and reports on responsiveness and quality of care in older populations.²⁰

Interpreting the measures

In interpreting the indicators of responsiveness, there is no clear cut-off between acceptable and unacceptable. Clearly, higher responsiveness levels and lower inequality measures are better. The literature shows that self-reported measures (e.g., responsiveness, quality of life, satisfaction) are right-skewed. This was illustrated in the WHO's raw survey results, which had 81% of respondents reporting in the highest two categories (range 52-96%) in the MCS Study and an average of 72% (range 38-92%) in the WHS. Therefore, the framework for interpreting the results on the WHS presented here adopts a benchmarking approach, comparing countries

with similar resource levels, based on the World Bank income classification of countries (see Appendix 3.1). The WHS classification of countries was incorporated for the European results – western European, and eastern European and former Soviet Union countries (Appendix 3.1).

Using this benchmarking approach and the analytical framework shown in Figure 3.1, we had some expectations of how the WHS results would look. We expected responsiveness to be higher in high resource settings because of the increased availability of human resources and better infrastructure. Human resources are the main conduit for the respect for person domains and, to some degree, prompt attention and choice. The higher the quality of the basic infrastructure in a country (e.g., better transport networks) the greater the impact on the domains of prompt attention and quality of basic amenities in health services.

We anticipate that there will be differences between responsiveness and general satisfaction measures for the same country although no direct comparison is drawn in this chapter. Measures of general satisfaction may respond to the contextual components described in Figure 3.1, but measures of responsiveness are based on actual experiences and would reflect the care process from the perspective of users.

WHS 2002 RESULTS

Sample statistics

The WHS 2002 was conducted in 70 countries, 68 of which reported back to WHO on their responsiveness data. Turkey did not complete the responsiveness section. The average interview completion response rate was 91% for all countries, ranging from 44% for Slovenia and up to 100% for as many as 22 countries. Note that the measure of survey response rates was interview completion rates – as mentioned, these rates may be as high as 100% as they express the number of persons who started and completed interviews as a percentage of the number of persons starting interviews. Sample sizes for ambulatory and inpatient care service averaged 1,530 and 609 respectively, across all countries. The wide range across countries (130–19,547 for ambulatory use in the last 12 months; 72–1,735 for inpatient use in the last 3 years) depended on both overall survey samples and different utilization rates across the different countries. Female participation in the overall survey sample averaged 56%, ranging from 41% (Spain) to 67% (Netherlands). The average age across all surveys was 43, ranging from 36 in Burkina Faso to 53 in Finland. Details on country-specific samples are provided in Appendix 3.2.

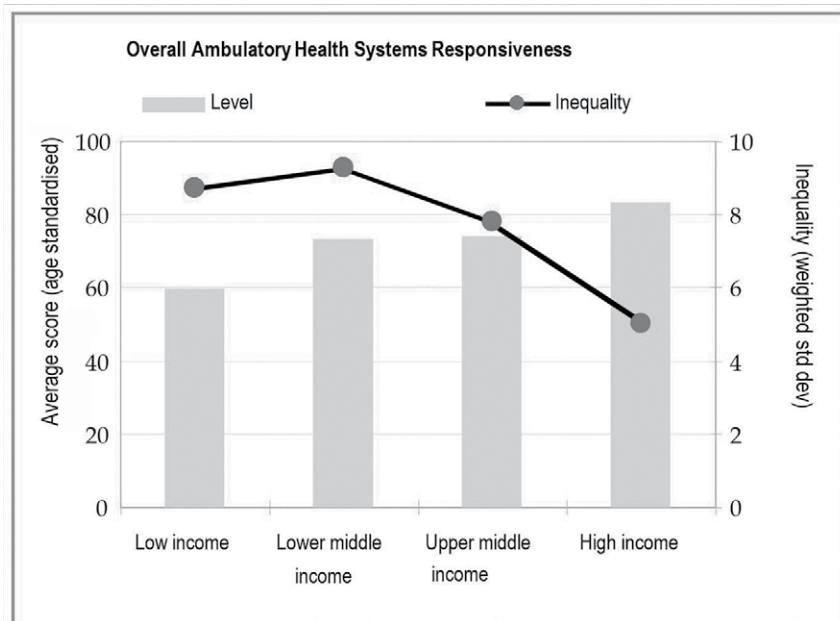


Figure 3.5 The level of inequalities in responsiveness by countries grouped according to World Bank income categories

Ambulatory care responsiveness

All countries

Overall results followed expected trends^v, with higher overall levels of responsiveness in higher-income countries as shown in Figure 3.5. Inequalities between lower- and middle-income countries changed slightly but, in general, large reductions in inequalities were only observed when moving from middle- to high-income countries.

Respondents from different country groupings consistently reported low responsiveness levels and high inequalities for the prompt attention domain. The dignity domain was consistently reported as high and its inequalities low. The overall gradient between country groupings as described in Figure 3.5 held for all domains. In other words, no domain was performing significantly better in a lower income grouping of countries than in the higher income grouping.

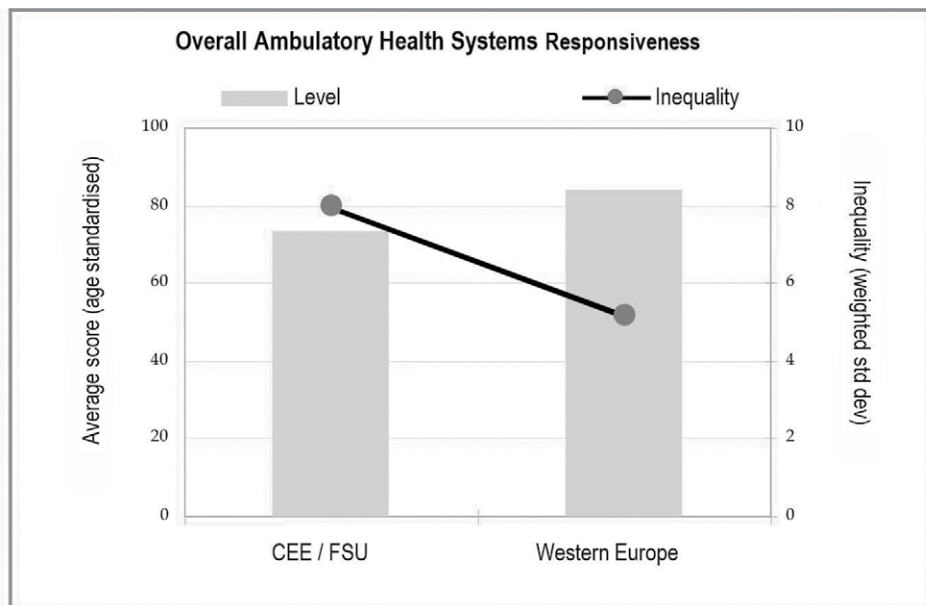


Figure 3.6 The level of inequalities in responsiveness by two groups of 25 European countries

European countries

Within Europe, western European countries shows notably higher mean levels of responsiveness and lower inequalities than the CEE/FSU countries (Figure 3.6). Responsiveness levels across all 25 European countries ranged from 56% in Russia to 92% in Austria (Figure 3.7). Inequalities ranged from 2.2 in Spain to 14.3 in Bosnia and Herzegovina. Strikingly, 9 of the 12 CEE/FSU countries had inequalities higher than the European average and only 4 of the 12 CEE/FSU countries had responsiveness levels greater than the average levels for Europe as a whole. By contrast, 12 of the 13 western European countries had responsiveness levels higher than the European average.

On average, responsiveness for all domains in western European countries was higher than in CEE/FSU countries. Differences were largest for the choice and autonomy domains. Prompt attention was the worst performing domain in Western Europe, while autonomy and prompt attention were the worst performing domains in CEE/FSU countries. Dignity was the best performing domain in both groups of countries, as found for the global average. Inequalities were higher for all domains in CEE/FSU countries. Both groups of countries had the highest inequal-

ities in the prompt attention domain. Inequalities were lowest in the communication domain in CEE/FSU countries and in the basic amenities and dignity domains in Western Europe.

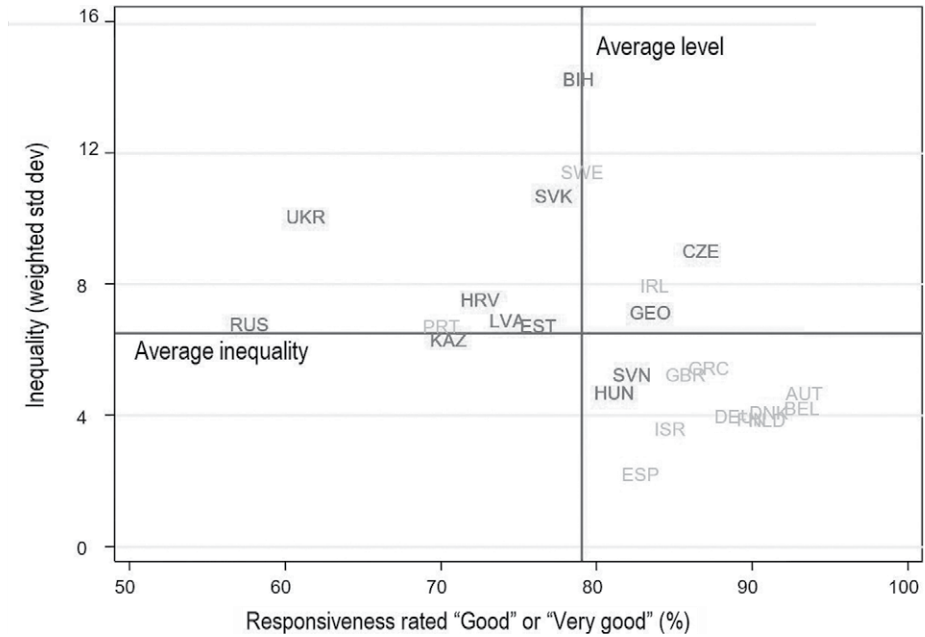


Figure 3.7 Inequalities in ambulatory responsiveness against levels for 25 European countries

Inpatient health services

All countries

The level of responsiveness for inpatient services increased across the four income groupings of countries (Figure 3.8)⁴. However, the pattern for inequalities was surprising. Unlike the trend seen in ambulatory care, inpatient inequalities reached a peak in upper middle income countries (greatest values in South Africa and Slovakia).

Responsiveness domains levels (except for autonomy and choice) increased across country groupings. Upper middle-income countries had lower levels of both domains compared to lower middle-income countries. In general, these domains were also the worst performing domains (compared with prompt attention for ambulatory services). The dignity domain performed best

in all groupings of countries, followed closely by social support. The spike in inequalities observed for upper middle-income countries seems to have arisen from sharply higher inequalities for the autonomy, basic amenities, and social support domains.

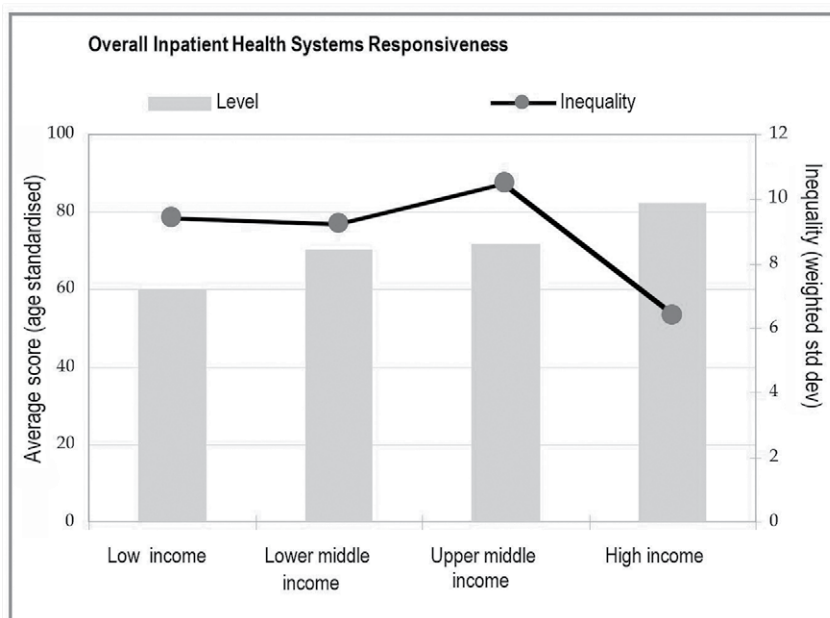


Figure 3.8 The level of inequality in responsiveness across World Bank income categories of countries

European countries

As for ambulatory services, responsiveness levels and inequalities in inpatient services differed between Western Europe and CEE/FSU countries (Figure 3.9). The average level of responsiveness levels across eleven CEE/FSU countries is 70% compared to 80% for 14 countries in Western Europe^{vii}. Inequalities were also higher in CEE/FSU countries.

Across all 25 European countries, responsiveness levels range from 51% in Ukraine to 90% in Luxembourg. Inequities range from a low of 3.4 in Austria to 18.9 in Slovakia. Ten of the 11 CEE/FSU countries (labelled in red in Figure 3.10) have responsiveness inequalities higher than that of the European average (for inequalities). Only 5 of the 11 CEE/FSU countries have responsive-

ness levels higher than the average level for Europe, whereas all 14 western European countries have a responsiveness level higher than the European average.

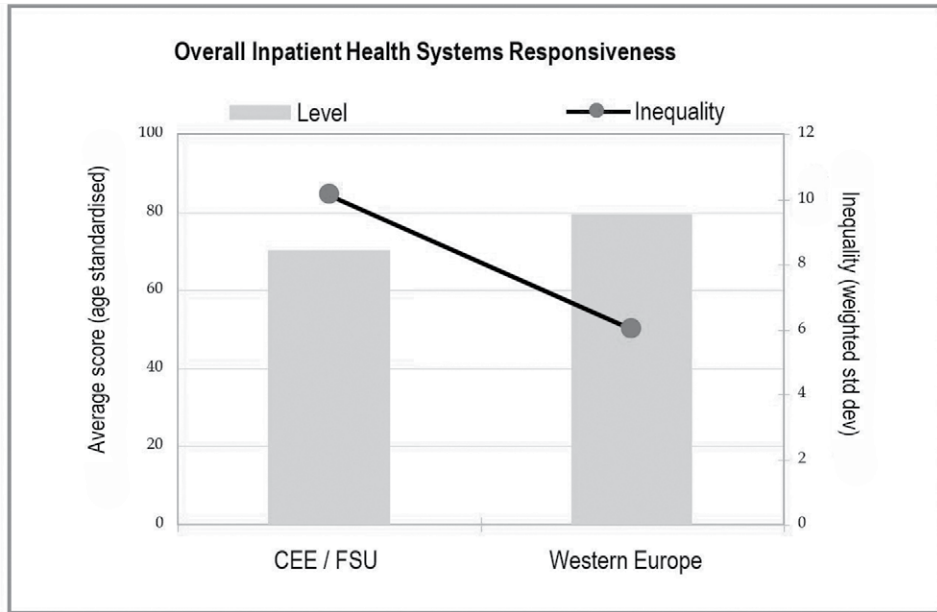


Figure 3.9 The level of inequalities in responsiveness by two groups of 25 European countries

As for ambulatory services, western European countries show higher levels for each of the 8 domains of inpatient services. Dignity was the best performing domain in CEE/FSU countries; in Western Europe both dignity and social support had the highest (similar) levels. Choice was the worst performing domain for both groups of countries.

Inequalities in all domains were higher for CEE/FSU countries; the highest inequality was seen in the prompt attention domain. In Western Europe, inequalities were highest in the domains of autonomy and confidentiality. In CEE/FSU countries the lowest inequalities were seen in the dignity domain while in Western Europe the lowest inequalities were seen in social support.

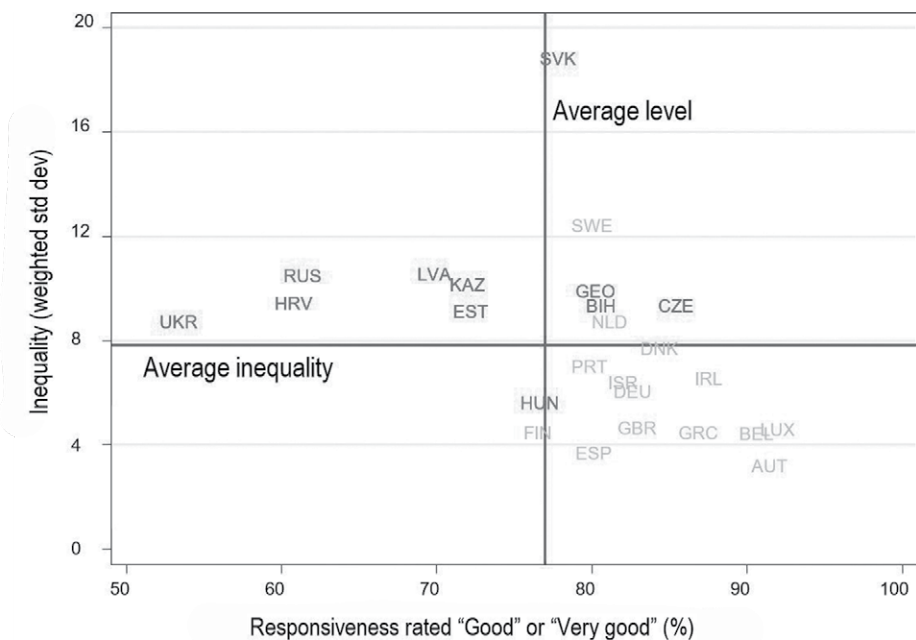


Figure 3.10 Responsiveness inequalities against average levels for 25 European countries

Responsiveness gradients within countries

Ambulatory health services

The values for the inequality indicator ranged between 5 and 10 for the different groups of countries. Figure 3.11 shows how these values translate into a gradient in responsiveness for different wealth or income quintiles within countries. Low- and middle-income countries showed a gradient but no gradient was seen in the high-income countries when averaged together.

In Europe, the CEE/FSU countries showed a gradient in the level of responsiveness across wealth quintiles with richer populations reporting better responsiveness (Figure 3.12). The gradient was nearly flat for western European countries.

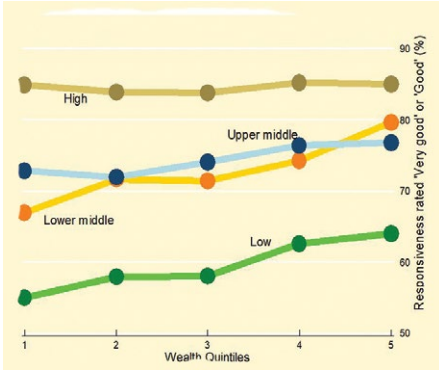


Figure 3.11 The gradient in responsiveness for population groups within countries by wealth quintiles

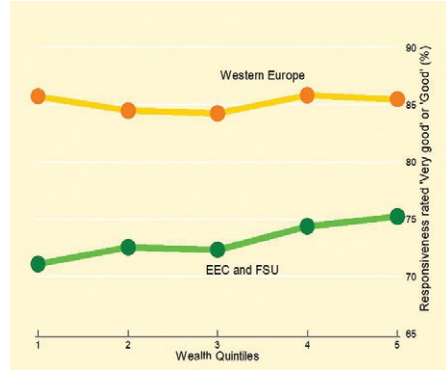


Figure 3.12 The gradient in responsiveness for population groups within countries in Europe by wealth quintiles

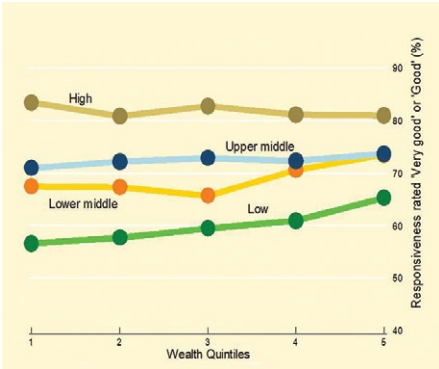


Figure 3.13 The gradient in responsiveness for population groups within countries by wealth quintiles

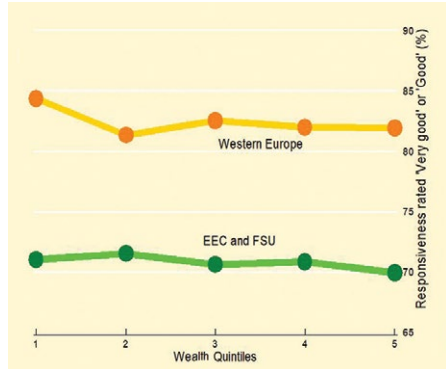


Figure 3.14 The gradient in responsiveness for population groups within countries in Europe by wealth quintile

Inpatient health services

The gradient in responsiveness for inpatient services is flatter than that observed for ambulatory services and most marked in low-income countries (Figure 3.13). Similarly, no gradient can be observed across wealth quintiles in the two groups of European countries. However, people in all quintiles in CEE/FSU countries clearly face worse levels of responsiveness than people in any quintile of Western Europe (Figure 3.14).

Health system characteristics and responsiveness

Figure 3.15 shows the rather obvious observation that factors such as resources in the health system provide a context to the process of care. It also shows the less obvious result that responsiveness affects the process of care, especially with respect to completion of treatment. We refer to this as coverage. With this understanding, we first explored the relationship between health expenditure and responsiveness in order to assess which domains might be more affected.

Second, we explored the relationship between responsiveness and indicators of completion of valid antenatal care as a means of understanding the relationship between responsiveness and coverage in general.

Keeping all other factors constant, well-resourced health system environments should be able to afford better quality care and receive better responsiveness ratings from users. Using a simple correlation for each responsiveness domain and keeping development contexts constant (by looking at correlations within World Bank country income groups), we observed whether higher health expenditures are associated with higher responsiveness and for which domains. Figure 3.15 lists the domains for which the correlations between total and government health expenditures and responsiveness are significant ($p=0.05$). In general, there is a positive association across many of the domains for most country income groupings, with the exception of lower middle-income countries. This indicates that increases in health expenditures in this grouping of countries are not being translated into improvements in patients' experiences of care, perhaps because absolute levels of expenditure are too low to create even a basic health system.

Where particular health needs require multiple contacts with the health system (e.g., chronic conditions or treatment protocols for TB or maternal care), the interaction between provider and user behaviours can influence utilization patterns. Under- or incorrect utilization can influence technical care and health outcomes^{viii 6}.

| | AMBULATORY | | INPATIENT | |
|-----------------------------|--|---|--|---|
| | Total health expenditure per capita | Government health expenditure per capita | Total health expenditure per capita | Government health expenditure per capita |
| Low-income (n, 19) | <ul style="list-style-type: none"> Higher levels for basic amenities, confidentiality Lower inequalities for dignity and autonomy | <ul style="list-style-type: none"> Higher levels for basic amenities, dignity, confidentiality Lower inequalities for dignity and basic amenities | <ul style="list-style-type: none"> Higher levels for basic amenities Lower inequalities for <u>all domains</u> except prompt attention | <ul style="list-style-type: none"> Higher levels for basic amenities Lower inequalities for dignity |
| Lower middle-income (n, 15) | <ul style="list-style-type: none"> None | <ul style="list-style-type: none"> None | <ul style="list-style-type: none"> None | <ul style="list-style-type: none"> Higher levels for dignity |
| High middle-income (n, 12) | <ul style="list-style-type: none"> Higher levels for communication, choice | <ul style="list-style-type: none"> Higher levels for dignity, communication, choice | <ul style="list-style-type: none"> Higher levels for choice, social support | <ul style="list-style-type: none"> Higher levels for prompt attention, choice, social support |
| High-income (n, 15) | <ul style="list-style-type: none"> Higher levels for communication, autonomy, choice, basic amenities Lower inequalities for basic amenities | <ul style="list-style-type: none"> Higher levels for <u>all domains</u> except confidentiality Lower inequalities for basic amenities | <ul style="list-style-type: none"> Higher levels for <u>all domains</u> except communication and confidentiality | <ul style="list-style-type: none"> Higher levels for <u>all domains</u> except confidentiality Lower inequalities for prompt attention, dignity, social support |

Figure 3.15 Correlations of average total health expenditure per capita and overall responsiveness for countries in different World Bank Income Categories

A few simple analyses of responsiveness and adherence-related data give a sense of the extent of validity in the WHS responsiveness results and how the acceptability and accessibility of services, as measured by responsiveness, can lead to adherence. Figure 3.16 shows a scatterplot of responsiveness and antenatal coverage rates. The latter rates were obtained from the WHS question which asked whether the respondent had completed four antenatal visits. Overall, a significant linear correlation was observed between the level of responsiveness and the percentage of respondents reporting that they had completed all four antenatal visits ($r=0.51$, $p=0.000$). The highest correlations were observed for the level of dignity (0.55), communication ($r=0.54$) and confidentiality (0.50). The responsiveness measure of inequality was less strongly correlated ($r=0.35$).

CONCLUSIONS

Empowering patients and equity in access are founding values that underpin the outlook for the new European health strategy. These values are expressed in the European White Paper, *Together for Health: A Strategic Approach for the EU 2008-2013*.³¹ Ensuring high responsiveness performance from health systems, both with respect to level and equity, is one key strategy to support these values and measuring responsiveness is one approach to keeping the issue high

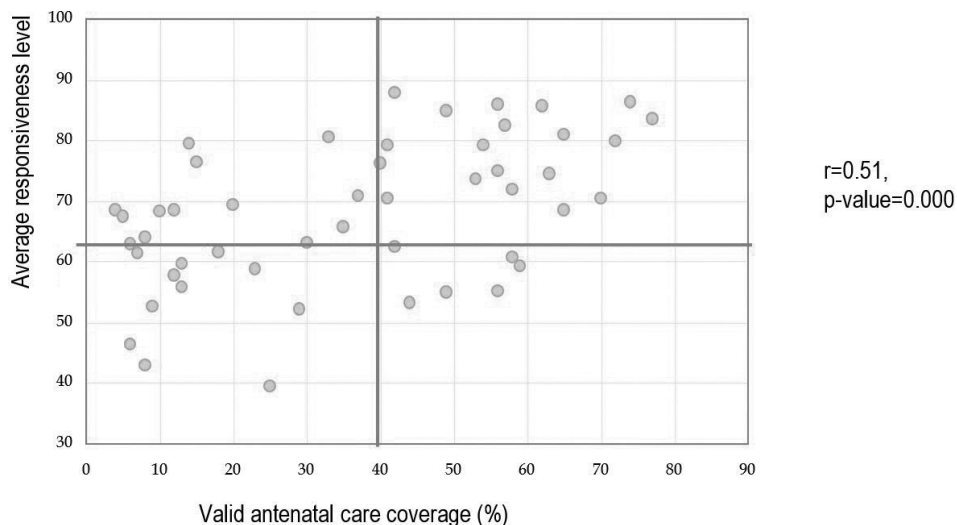


Figure 3.16 Responsiveness and antenatal coverage

on the health systems performance agenda. The analyses for this chapter used inequalities in responsiveness across income groups as a proxy for inequities in responsiveness. The discussion below refers to these two aspects of responsiveness.

Common concerns

A wide array of results on health system responsiveness has been presented in this chapter. Health systems across the world show some common strengths and failings. Nurses' and doctors' respectful treatment of users is encapsulated in the responsiveness domain – dignity. This is a relative strength in comparison to systemic issues such as prompt attention, involvement in decision making (autonomy), or choice (continuity of provider).

Our analysis has generally confirmed the hypothesis of a positive relationship between a country's level of development (represented by national income) and the responsiveness of its health system (as is observed for health outcomes). However, while there is a linear relationship between the income level in a country and the average level of responsiveness, dramatic reductions in responsiveness inequalities are only observed in the high-income country category. This observation was true for both inpatient and ambulatory care.

Elevated levels of health expenditures in a system are no guarantee that a system's responsiveness has improved. For lower middle-income countries no gains in responsiveness are observed for increases in health expenditures, probably due to inadequate general funding. Increased health expenditure (particularly in the public sector) for the other country groupings does yield gains in the overall responsiveness level and equality, but usually in some specific domains. On the other hand, lower responsiveness is associated with lower coverage, and inequalities in responsiveness are associated with greater inequity in access, regardless of development setting. Hence, explicit steps need to be taken to build good levels of responsiveness performance into all systems.

The European analysis showed that there are substantial differences in mean levels and within-country inequalities between western European, and CEE/FSU countries. Average responsiveness levels are higher in western European (85%) countries than in CEE/FSU (73%) countries. In both groups of countries, ambulatory services had the highest levels of responsiveness for dignity and the highest inequalities for prompt attention. In inpatient services, levels of dignity were highest in both country groupings, but prompt attention inequalities were highest in CEE/FSU countries and autonomy and confidentiality inequalities were highest in Western Europe.

Implementing change

Enhancing communication in the health system provides a potential entry point for improving responsiveness. Clear communication is associated with dignity, better involvement in decision-making and, in addition, supports better coverage or access. It is also an attribute that is highly valued by most societies. In the European context, it is interesting to note that CEE/FSU countries place special importance on communication.³²

As shown here, responsiveness appears to be complementary or contributory to ensuring equity in access (to the technical quality of care). This is in keeping with the Aday and Andersen⁷ framework and with Donabedian² who introduced the concept of the quality of health care and satisfaction with the care received as a valid component for achieving high technical quality of care and high rates of access to care. Inequities in access will result if the process of care systematically dissuades some groups from either initiating or continuing use of services to obtain the maximum benefit from the intervention. It is critical to deliver health interventions effectively and ensure compliance in primary care where a large majority of the population receives preventive and promotive health interventions. This is likely to become an increasing concern with the global epidemiological transition from infectious to chronic diseases. Therefore, primary care providers need to be aware of their critical role in patient communication and treating individuals with respect.

Responsiveness measurement and future research

The psychometric properties of the responsiveness questions show resilience across different countries and settings and indicate that the responsiveness surveys (when reported as raw data) have face validity. The WHS managed to improve on the MCS Study questions in several ways and provides a useful starting tool for countries embarking on routine assessments of responsiveness.

Some key aspects of responsiveness still need to be researched further. In particular, while theoretically complementary, further investigation could benefit empirical research on the potential trade-offs between health (through investments in improved technical applications) and non-health (through better responsiveness) outcomes.

A second key area relates to gaining a better understanding of how responsiveness and responsiveness inequities may act as indicators of inequities in access or unmet need in the population and what measures can be taken to improve responsiveness in the light of this relationship.

A third key area relates to the self-reported nature of the responsiveness instrument. Self-reported data may be prone to measurement error^{33,34} where bias results from groups of respondents (for example defined by socio-economic characteristics) varying systematically in their reporting of a fixed level of the measurement construct. The degree of comparability of self-reported survey data across individuals, socio-economic groups or populations has been debated extensively, usually with regard to health status measures.^{35,36}

Similar concerns apply to self-reported data on health systems responsiveness where the characteristics of the systems and cultural norms regarding the use and experiences of public services are likely to predominate. The method of anchoring vignettes has been promoted as a means for controlling for systematic differences in preferences and norms when responding to survey questions.³⁷ Vignettes represent hypothetical descriptions of fixed levels of a construct (such as responsiveness) and individuals are asked to evaluate these in the same way that they are asked to evaluate their own experiences of the health system. The vignettes provide a source of external variation from which information on systematic reporting behaviour can be obtained. To date, little use has been made of the vignette data within the WHS³⁸ and these offer a valuable area for future research.

Prospects for measuring responsiveness

Non-health outcomes are gaining increasing attention as valid measures of performance and quality. These require some feedback on what happens when users make contact with health care systems that can be easily compared across countries. Routine surveys on responsiveness are by no means a substitute for other forms of participation but, within the theme of patient empowerment, can provide opportunities for users' voices to be heard in health care systems.

Responsiveness measurement (as opposed to broader patient satisfaction measurement) is increasingly recognized as an appropriate approach for informing health system policy. Work by the Picker Institute³⁹ and the AHRQ⁴⁰; the future work envisaged by the OECD⁴¹; and the broader analytical literature have built this case very satisfactorily. The work of the last decade has provided a solid base and an opportunity for individual countries to introduce measures of responsiveness into their health-policy information systems in the short and medium term.

REFERENCES

1. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB. Health system responsiveness: concepts, domains and measurement. In: Murray CJL, Evans DB, editors. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003. p. 573–96.
2. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
3. World Health Organization. Health and human rights and equity working group draft glossary. 2005.
4. World Health Organization. Health systems: improving performance. The world health report 2000. Geneva: WHO; 2000.
5. Murray CJL, Frenk J. A framework for assessing the performance of health systems. Bull World Health Organ. 2000;78(6):717–31.
6. Donabedian A. Aspects of medical care administration. Cambridge, MA: Harvard University Press; 1973.
7. Aday L, Andersen R. A framework for the study of access to medical care. Health Serv Res. 1974;9:208–20.
8. Tanahashi T. Health service coverage and its evaluation. Bull World Health Organ. 1978;56(2):295–303.
9. Andersen R. Revisiting the behavioural model and access to medical care: does it matter? J Health Soc Behav. 1995;36(1):1–10.
10. Bradley E, McGraw S, Curry L, Buckser A, King K, Kasl S, et al. Expanding the Andersen model: the role of psychosocial factors in long-term care use. Health Serv Res. 2002;37(5):1221–42.
11. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. Social determinants of health discussion paper 2. 2nd 2010 ed. Geneva: WHO; 2007.
12. Gilson L, Doherty J, Loewenson R, Francis V. Challenging inequity through health systems: final report. Knowledge Network on Health Systems (WHO Commission on Social Determinants of Health). Geneva: World Health Organization; 2007. [online: http://www.who.int/social_determinants/resources/csdh_media/hskn_final_2007_en.pdf]
13. Shengelia B, Tandon A, Adams O, Murray CJL. Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy. Soc Sci Med. 2005;61(1):97–109.
14. Ware JE, Hays RD. Methods for measuring patient satisfaction with specific medical encounters. Med Care. 1988;26(4):393–402.
15. World Health Organization, EQUINET-REACH TRUST Malawi, TARSC Zimbabwe. Monitoring equity in access to AIDS treatment programmes: a review of concepts, models, methods and indicators. Geneva: WHO; 2010.
16. World Health Organization. Technical consultation on concepts and methods for measuring the responsiveness of health systems. In: Murray CJL, Evans DB, editors. Health systems performance assessment: debates, methods and empiricism. Geneva: WHO; 2003. p.115–24.

17. Blendon R, Schoen C, DesRoches C, Osborn R, Zaper K. Common concerns amid diverse systems: health care experiences in five countries. *Health Affair*. 2003;22(3):106–21.
18. de Silva A. A framework for measuring responsiveness. Global programme on evidence for health policy discussion paper no. 32. Geneva: World Health Organization; 2000.
19. Üstün TB, Chatterji S, Villanueva M, Bendib L, Celik C, Sadana R, Valentine NB, Ortiz J, Tandon A, Salomon J, Cao Y, Wan Jun X, Özaltın E, Mathers C, Murray CJL. WHO Multi-Country Survey Study on Health and Responsiveness 2000-2001. Global programme on evidence for health policy. Discussion paper no. 37. Geneva: World Health Organization; 2001.
20. Valentine NB, Bonsel GJ, Murray CJL. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Qual Life Res*. 2007;16(7):1107–25.
21. Üstün TB, Chatterji S, Mechbal A, Murray CJL, WHS Collaborating Groups. The World Health Surveys. Chapter 58. In: Murray CJL, Evans DB, editors. *Health Systems Performance Assessment: debates, methods and empiricism*. Geneva: World Health Organization; 2003. p.797–808.
22. Valentine NB, Lavalley R, Liu B, Bonsel GJ, Murray CJL. Classical psychometric assessment of the responsiveness instrument in the WHO Multi-Country Survey Study on Health and Responsiveness 2000–2001. In: Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism*. Geneva: WHO; 2003. p. 597–630.
23. World Health Organization. The health systems responsiveness analytical guidelines for surveys in the Multi-Country Survey Study 2005 [Internet]. Geneva: World Health Organization; 2005 (December). Available from: http://www.who.int/responsiveness/papers/MCSS_Analytical_Guidelines.pdf?ua=1
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
25. Üstün TB, Chatterji S, Mechbal A, Murray CJL. Quality assurance in surveys: standards, guidelines and procedures. *Household surveys in developing and transition countries: design, implementation and analysis*. New York: United Nations; 2005.
26. Labarere J, Francois P, Auquier P, Robert C, Fourny M. Development of a French inpatient satisfaction questionnaire. *Int J Qual Health C*. 2001;13:99–108.
27. Steine S, Finset A, Laerum E. A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. *Fam Pract*. 2001;18(4):410–9.
28. Nunnally JC, Bernstein IH. *Psychometric theory*, 3rd ed ed. New York: McGraw-Hill; 1994.
29. Hall J, Feldstein M, Fretwell M, Rowe J, Epstein A. Older patients' health status and satisfaction with medical care in an HMO population. *Med Care*. 1990;28:261–70.
30. Harper S, Lynch J. Measuring health inequalities. In: Oakes J, Kaufman J, editors. *Methods in social epidemiology*. San Francisco: John Wiley & sons; 2006. p. 134–68.
31. Commission of the European Communities. Together for health: a strategic approach for the EU 2008–2013. White paper. Com(2007) 630 final (Oct. 23, 2007). [online: http://ec.europa.eu/health/ph_overview/Documents/strategy_wp_en.pdf]

Chapter 3. Health System Responsiveness by N Valentine

32. Valentine NB, Darby C, Bonsel GJ. Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of "health systems responsiveness" in 41 countries. *Soc Sci Med*. 2008;66(9):1939–50.
33. Groot W. Adaptation and scale of reference bias in self-assessments of quality of life. *J Health Econ*. 2000;19:403–20.
34. Murray CJL, Kawabata K, Valentine NB. People's experience versus people's expectations. *Health Affairs*. 2001;20(3):41–3.
35. Bago d'Uva T, van Doorslaer E, Lindeboom M. Does reporting heterogeneity bias the measurement of health disparities? *Health Econ*. 2008;17:351–75.
36. Lindeboom M, van Doorslaer E. Cut-point shift and index shift in self-reported health. *J Health Econ*. 2004;23:1083–99.
37. Salomon J, Tandon A, Murray CJL. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. *Brit Med J*. 2004;328(7434):258.
38. Rice N, Robone S, Smith P. The measurement and comparison of health system responsiveness. Health econometrics and data group (HEDG) working paper 08/05. York: University of York; 2008.
39. Picker Institute. The Picker Institute implementation manual. Boston, MA: Picker Institute; 1999.
40. AHRQ. CAHPS 2.0 survey and reporting kit. Rockville: Agency for Healthcare Research and Quality; 1999.
41. Garratt A, Solheim E, Danielsen K. National and cross-national surveys of patient experiences: a structured review. Report no. 7-2008. Oslo: Norwegian Knowledge Centre for the Health Services; 2008.

ENDNOTES

ⁱ This definition was suggested by Elias Mossialos, who commented on the draft chapter.

ⁱⁱ This sort of analysis makes no sense with reference to countries where the short version questionnaire has been implemented, since only one item was present in each domain for these countries.

ⁱⁱⁱ The indicator from Harper and Lynch (2006)³⁰ was further modified by Dr. Ahmad Hosseinpoor (WHO/IER). The title of the paper is 'Global inequalities in life expectancy among men and women' (tentative). It was finally published as 'International shortfall inequality in life expectancy in women and in men: 1950-2010' (Hosseinpoor AR, Harper S, Lee JL, Lynch J, Mathers C, Abou-Zahr C. *Bull World Health Organ.* 2012; 90:588-94).

^{iv} The formula:
$$\frac{\sum_{j=1}^J N_j |y_j - \mu|}{N}$$
 ; y_j : the rate in group j , μ : the rate in reference group, N_j : population size of each group, N : total population

^v The following countries were not included as they did not record an ambulatory section: Australia, France, Norway and Swaziland. In addition, the following countries were dropped from the analysis as their datasets did not have (minimum) sufficient observations for each quintile (30 or more): Italy, Luxembourg, Mali and Senegal.

^{vi} The following countries were not included in the analysis: (1) Australia, France, and Norway for lack of data on assets for construction of wealth index; (2) Swaziland for too few observations in the ambulatory section. Also, the following countries were dropped from the analysis as their datasets did not have (minimum) sufficient observations for each quintile: Ethiopia, Italy, Mali, Senegal, and Slovenia.

^{vii} Italy and Slovenia were omitted from the inpatient services analysis as their datasets did not have the minimum number of observations for reliable results.

^{viii} This assumes that, when applied technically correctly, health interventions have a positive impact on health.

Appendix 3.1 Groupings of World Health Survey countries

WHS countries grouped by World Bank income categories

| | |
|---|---|
| <p>Low income Bangladesh, Burkina Faso, Chad, Comoros, Congo, Cote d'Ivoire, Ethiopia, Ghana, India, Kenya, Lao, Malawi, Mali, Mauritania, Myanmar, Nepal, Pakistan, Senegal, Vietnam, Zambia, Zimbabwe</p> | <p>Lower middle income Brazil, China, Ecuador, Georgia, Guatemala, Kazakhstan, Morocco, Namibia, Paraguay, Philippines, Sri Lanka, Tunisia, Ukraine, Bosnia and Herzegovina, Dominican Republic</p> |
| <p>Higher middle income Croatia, Czech, Estonia, Hungary, Latvia, Malaysia, Mauritius, Mexico, Russia, Slovakia, South Africa, Uruguay</p> | <p>High income Austria, Belgium, Denmark, Finland, Germany, Greece, Ireland, Israel, Italy, Luxembourg, Netherlands, Portugal, Slovenia, Spain, Sweden, United Arab Emirates (UAE), United Kingdom (UK)</p> |

WHS countries in Europe

| | |
|--|---|
| <p>Central and Eastern Europe, and Former Soviet Union (CEE/FSU)</p> | <p>Western Europe</p> |
| <p>Bosnia and Herzegovina, Croatia, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Latvia, Russia, Slovakia, Slovenia, Ukraine</p> | <p>Austria, Belgium, Denmark, Finland, Germany, Greece, Ireland, Israel, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom (UK)</p> |

Appendix 3.2 WHS 2002 sample descriptive statistics

| Country | Response rate/ interview completion (percentage) | Users of ambulatory services in last 12 months | Users of inpatient services in the last 3 years | Percentage female | Average age (years) | Percentage high school or more educated | Percentage in good or very good health |
|----------------------------|--|---|--|----------------------|---------------------------|--|---|
| Low income | | | | | | | |
| Bangladesh | 85 | 4020 | 777 | 53 | 39 | 8 | 44 |
| Burkina Faso | 96 | 1199 | 589 | 53 | 36 | 3 | 70 |
| Chad | 92 | 423 | 371 | 53 | 37 | 3 | 58 |
| Comoros | 95 | 526 | 374 | 55 | 42 | 5 | 54 |
| Congo | 79 | 381 | 288 | 53 | 36 | 18 | 56 |
| Cote d'Ivoire | 97 | 765 | 305 | 43 | 36 | 13 | 60 |
| Ethiopia | 96 | 1779 | 224 | 52 | 37 | 3 | 75 |
| Ghana | 70 | 1567 | 677 | 55 | 41 | 4 | 72 |
| India | 93 | 5003 | 1735 | 51 | 39 | 21 | 58 |
| Kenya | 82 | 2228 | 803 | 58 | 38 | 21 | 66 |
| Lao | 98 | 735 | 570 | 53 | 38 | 10 | 78 |
| Malawi | 93 | 2423 | 1236 | 58 | 36 | 1 | 79 |
| Mali | 79 | 130 | 104 | 43 | 42 | 3 | 70 |
| Mauritania | 98 | 552 | 469 | 61 | 39 | 10 | 69 |
| Myanmar | 97 | 1667 | 320 | 57 | 41 | 9 | 79 |
| Nepal | 98 | 3279 | 1141 | 57 | 39 | 5 | 62 |
| Pakistan | 93 | 3727 | 913 | 44 | 37 | 14 | 75 |
| Senegal | 88 | 222 | 182 | 48 | 38 | 8 | 58 |
| Vietnam | 84 | 1541 | 548 | 54 | 40 | 24 | 51 |
| Zambia | 88 | 2188 | 764 | 55 | 36 | 5 | 72 |
| Zimbabwe | 94 | 1660 | 649 | 64 | 37 | 5 | 52 |
| Lower middle income | | | | | | | |
| Bosnia and Herzegovina | 94 | 394 | 259 | 58 | 47 | 8 | 58 |
| Brazil | 100 | 2341 | 1244 | 56 | 42 | 28 | 53 |
| China | 100 | 1435 | 423 | 51 | 45 | 28 | 62 |
| Dominican Republic | 74 | 1315 | 1508 | 54 | 42 | 5 | 56 |
| Ecuador | 77 | 1372 | 592 | 56 | 41 | 13 | 57 |
| Georgia | 92 | 763 | 227 | 58 | 49 | 88 | 38 |
| Guatemala | 98 | 2063 | 978 | 62 | 40 | 12 | 53 |
| Kazakhstan | 100 | 2331 | 803 | 66 | 41 | 96 | 48 |
| Morocco | 79 | 2211 | 800 | 59 | 41 | 14 | 41 |
| Namibia | 91 | 650 | 862 | 59 | 38 | 4 | 72 |
| Paraguay | 97 | 2414 | 1096 | 54 | 40 | 12 | 70 |
| Philippines | 100 | 2625 | 906 | 52 | 39 | 16 | 60 |

Appendix 3.2 continued

| Country | Response rate/ interview completion (percentage) | Users of ambulatory services in last 12 months | Users of inpatient services in the last 3 years | Percentage female | Average age (years) | Percentage high school or more educated | Percentage in good or very good health |
|----------------------------|---|--|---|----------------------|---------------------------|--|---|
| Sri Lanka | 99 | 2268 | 1697 | 53 | 41 | 21 | 72 |
| Tunisia | 96 | 2352 | 816 | 53 | 42 | 28 | 62 |
| Ukraine | 99 | 735 | 580 | 64 | 48 | 87 | 27 |
| Upper middle income | | | | | | | |
| Croatia | 100 | 465 | 259 | 59 | 52 | 16 | 51 |
| Czech Republic | 49 | 411 | 302 | 55 | 48 | 47 | 55 |
| Estonia | 99 | 395 | 289 | 64 | 50 | 74 | 36 |
| Hungary | 100 | 453 | 489 | 58 | 49 | 63 | 51 |
| Latvia | 92 | 283 | 293 | 67 | 51 | 34 | 33 |
| Malaysia | 80 | 1943 | 1329 | 56 | 41 | 42 | 78 |
| Mauritius | 88 | 1702 | 1180 | 52 | 42 | 13 | 65 |
| Mexico | 97 | 19457 | 1440 | 55 | 42 | 23 | 67 |
| Russia | 100 | 1794 | 1019 | 64 | 51 | 61 | 31 |
| Slovakia | 99 | 897 | 355 | 62 | 39 | 71 | 66 |
| South Africa | 89 | 384 | 384 | 53 | 38 | 34 | 73 |
| Uruguay | 100 | 1029 | 536 | 51 | 46 | 30 | 79 |
| High income | | | | | | | |
| Austria | 100 | 184 | 351 | 62 | 45 | 26 | 77 |
| Belgium | 100 | 298 | 299 | 56 | 45 | 64 | 74 |
| Denmark | 100 | 316 | 194 | 53 | 51 | 52 | 79 |
| Finland | 100 | 464 | 345 | 55 | 53 | 58 | 55 |
| Germany | 100 | 428 | 401 | 60 | 50 | 23 | 65 |
| Greece | 100 | 433 | 272 | 50 | 51 | 47 | 67 |
| Ireland | 100 | 239 | 214 | 55 | 44 | 19 | 82 |
| Israel | 57 | 521 | 412 | 57 | 45 | 85 | 76 |
| Italy | 100 | 541 | 232 | 57 | 48 | 51 | 63 |
| Luxembourg | 100 | 135 | 237 | 52 | 45 | 43 | 73 |
| Netherlands | 100 | 624 | 192 | 67 | 44 | 83 | 76 |
| Portugal | 100 | 510 | 212 | 62 | 50 | 20 | 39 |
| Slovenia | 44 | 284 | 72 | 53 | 47 | 52 | 58 |
| Spain | 53 | 2863 | 1601 | 41 | 53 | 31 | 64 |
| Sweden | 100 | 300 | 266 | 58 | 51 | 70 | 62 |
| UAE | 100 | 453 | 239 | 48 | 37 | 65 | 86 |
| UK | 100 | 369 | 344 | 63 | 50 | 46 | 68 |

CHAPTER 4

Health systems' responsiveness and reporting behaviour: multilevel analysis of the influence of individual-level factors in 64 countries



N.B. Valentine, E.Verdes-Tennant, G.J. Bonsel
Social Science & Medicine. 2015; 138:152-60

ABSTRACT

Health systems responsiveness encompasses attributes of health system encounters valued by people and measured from the user's perspective in 8 domains: dignity, autonomy, confidentiality, communication, prompt attention, social support, quality of basic amenities and choice. The literature advocates for adjusting responsiveness measures for reporting behaviour heterogeneity, which refers to differential use of the response scale by survey respondents. Reporting behaviour heterogeneity between individual respondents compromises comparability between countries and population subgroups. It can be studied through analysing responses to pre-defined vignettes – hypothetical scenarios recounting a third person's experience in a health care setting. This paper describes the first comprehensive approach to studying reporting behaviour heterogeneity using vignettes. Individual-level variables affecting reporting behaviour are grouped into three categories: (1) sociodemographic, (2) health-related and (3) health value system. We use cross-sectional data from 150,000 respondents in 64 countries from the World Health Organization's World Health Survey (2002-03). Our approach classifies effect patterns for the scale as a whole, in terms of strength and in relation to the domains. For the final 8 variables selected (sex; age; education; marital status; use of inpatient services; perceived health (own); caring for close family or friends with a chronic illness; the importance of responsiveness), the strongest effects were present for education, health, caring for friends or relatives with chronic health conditions, and the importance of responsiveness. Patterns of scale elongation or contraction were more common than uniform scale shifts and were usually constant for a particular factor across domains. The dependency of individual-level reporting behaviour heterogeneity on country is greatest for prompt attention, quality of basic amenities and confidentiality domains.

INTRODUCTION

Progress in the health sector can be evaluated by changes in population health. But different monitoring frameworks also acknowledge the importance of other indicators. For example, in 2000 the World Health Organization (WHO) estimated additional measures of people-centredness of a health system, termed responsiveness, and fairness in financial contributions. The results were widely debated at the time. However, the underlying framework was generally accepted.¹ Recent global policy debates have again shifted the focus onto actionable indicators of health systems performance in addition to health, such as coverage, out-of-pocket expenditures and people-centredness of the health system.²

Responsiveness encompasses the acceptability of service provision with reference to the way users are treated and the environment within which they treated. Encompassing a range of issues that are important to people, it has also been referred to as responding to a population's "legitimate expectations" regarding the characteristics of an acceptable service³ (page 719). Responsiveness contributes to satisfaction, well-being, and human dignity.⁴ It also has instrumental value for achieving other objectives such as improving treatment success rates.⁵⁻⁷

The measurement of constructs like responsiveness commonly requires surveys with questions for individuals with recent health service experiences (e.g., Measure Demographic and Household Survey (DHS) Antenatal Client Exit Interviews: <http://www.measuredhs.com/>; Consumer Assessment of Health Providers and Systems (CAHPS)). Aggregation of individuals' responses to these questions provides a user-based assessment of quality. The World Health Organization has a large publicly available dataset on responsiveness from the World Health Survey (WHS) (2002-2003). The WHS responsiveness questions covered 8 'domains', named as follows (the short name appears in parenthesis where applicable): involvement in decision-making ('autonomy'), respectful treatment ('dignity'), clear communication ('communication'), confidentiality of personal information ('confidentiality'), choice of provider ('choice'), prompt attention, quality of basic amenities, access to family and community support ('social support'). The questions on domain performance use ordinal response scales with the verbal response categories, "very good" (1), "good" (2), "moderate" (3), "bad" (4), "very bad" (5).

A central challenge to comparing responsiveness survey results is known as 'differential scale use' or 'reporting behaviour heterogeneity'.⁸ It refers to the differential use of the response scale by respondents, unrelated to the object measured. Reporting behaviour heterogeneity consists of a random and systematic component. If the non-random component is related to the comparator of interest (e.g., culture, socioeconomic status), it compromises comparability across countries and within countries.⁹ Such measurement error is not unique to responsiveness. It is also

found in other subjective measures of latent constructs, like happiness, well-being or self-reported health. If significant, adjustment or rescaling of people's evaluations may be required for comparisons within and across countries.

Differential scale use arises for different reasons. Social norms may cause some populations to avoid extreme expressions of approval or disapproval. Better education may increase comprehension and judgement abilities and the corresponding differentiation of the respondent between the verbal response categories ("very good" to "very bad"). Several studies have found that education is an important determinant of reporting behaviour heterogeneity.^{9,10} Other factors that affect judgement may relate to familiarity with health problems and services^{11,12} or age¹². Differential use of the response scale may also arise from differing 'expectations' associated with social status¹³ as demarcated by income or ethnicity.¹⁴ Individuals accustomed to poorer quality experiences could have lower expectations, causing their judgements to be less harsh, even when receiving worse quality services. Another mechanism affecting scale use is related to the importance an individual assigns to an attribute of the care process. Importance fixes attitudes more firmly, creating a stronger sense of expected norms, which has been shown to change the range of scale used.^{15,16}

There are different ways of addressing reporting behaviour heterogeneity. Stratification of data by particular social groups without standardization tends to ignore the problem related to scale use but can focus on views of sicker patients as "bellwethers for how well health care systems are working" ¹⁷(page 106). Standardizing stratified results by personal characteristics found affecting patient assessments of their own experiences in regression analyses is an implicit adjustment procedure. This has also been referred to as patient-mix adjustment. It has been used in reporting by AHRQ on CAHPS surveys¹² and as part of the Picker Survey methodology.¹⁸ But again, even if patient-mix adjustment is used, it can purge valid disparities or inequities in health care responsiveness.

An increasingly common approach to adjustment, is to characterize reporting behaviour based on respondents' answers to a separate set of questions from those concerning the respondent's experiences. These are called vignette questions. Whereas performance or assessment questions ask users how they evaluate the health systems responsiveness during their own experiences, vignettes describe hypothetical, reference health care situations that a third person is experiencing and vignette questions request users to evaluate these situations. Vignette questions use the same response scale as performance questions and provide researchers with specific information, distinguishing scale use from actual patient experiences. This information can be used to adjust the survey respondent's rating of their own experience through standardization or other techniques (parametric or non-parametric). Vignette-based adjustment proce-

dures have early 20th century roots: in 1948 the physicist S. L. Anderson introduced a technique for scoring slubs in a wool yarn on a five-point scale, using four reference specimens to adjust the scores of judges for systematic scoring heterogeneity.¹⁹ In a similar vein the WHO surveys of the MCS and WHS introduced the equivalent of reference specimens known as ‘anchoring vignettes’.

The WHS vignettes have been successfully tested for eliciting a common understanding across respondents (vignette equivalence).^{9,20} The data have been used to publish papers quite widely^{21,22}, which include studies on their psychometric properties.²³

Vignettes have subsequently also been used in non-WHO surveys (e.g., SHARE 2006-07, Wisconsin Longitudinal Study) and in several specialized procedures that adjust for systematic reporting of behaviour heterogeneity where their use was found to improve comparability in both WHO and non-WHO surveys.^{24,25}

This paper has the explicit goal of analysing vignettes from the World Health Survey in order to characterize how a broad range of individual-level factors affect respondents’ reporting behaviour. In general, previous studies reporting adjustment procedures with anchoring vignettes have taken account of a limited range of individual-level factors from the potential array of factors to be considered. Their usual focus has produced results with fewer applications for practitioners wanting to compare responsiveness within countries to further quality improvements. The data analysed in this paper cover a wider range of individual-level factors than used previously in a single model, for all responsiveness domains, while retaining a model structure that allows for country-level effects as done elsewhere. It aims to contribute to a general adjustment framework, and associated reporting standards within countries by characterizing different aspects of reporting behaviour heterogeneity, including how to describe the observed systematic influences on the use of reporting scale. If our study reveals that individual-level variables are important, then adjustments of responsiveness results within countries is needed to ensure local comparability. This knowledge and explicit characterization of reporting behaviour patterns, will also improve the validity of between country comparisons of responsiveness.

METHODS

Survey organization and questionnaire

The World Health Survey was a household-based survey administered in 71 countries in 2002/03 with datasets finalized in 2004/5. The surveys used stratified, multiple cluster, designs. Ethical approval was obtained from an independent ethics review conducted by the Harvard School of Public Health’s Institutional Review Board.²⁶ Questionnaires were translated into

the principal languages of the participating countries, with back-translation of all key terms. Pre-testing and cognitive interviewers were used to finalize the questionnaire and its translation, and to smooth out operational problems before the survey launch. Surveys were face-to-face surveys except for in Israel, Luxembourg and Norway (telephone). Data collection was carried out by multiple experienced national or multinational contractors (e.g., GALLUP), using detailed protocols. The questionnaire comprised modules on socio-demographic background, social capital, own health, own health care utilization, own responsiveness experience, and health and responsiveness vignettes. (For details, see: <http://www.who.int/healthinfo/survey/en/>)

Dependent variable: responsiveness domain scale

Respondents with at least one recent health service experience (inpatient, the past 5 years; outpatient, the past 12 months) were invited to answer questions evaluating responsiveness vignettes on a scale from “very good” to “very bad”. A total of 40 different vignettes, with one vignettes question per vignette, were used over 4 vignette module rotations, effectively distributing 10 vignettes covering 2 domains to each quarter of a survey’s respondents. Evaluating all 40 vignettes was considered too demanding for a single respondent although it was necessary for each respondent to evaluate all 5 vignettes in a particular domain, in order to cover the response scale. The rotations of domains were as follows: set A combined prompt attention and dignity vignettes; set B - communication and quality of basic amenities; set C - confidentiality and choice; and set D - access to social support and autonomy. As an example, the first vignette and associated vignette question in the domain of dignity, from rotation set A, is shown in Figure 4.1. Previous studies of the ranking of these vignettes across all 8 domains have demonstrated their validity in relation to distinguishing between different levels of service standards (from best state to worst state), as measured by high inter-individual Spearman’s rank ordered coefficients.²⁰ This implies that respondents understood the vignettes in the same way, and hence the data did not violate the essential ‘vignette equivalence’ assumption.

Study data

This paper uses pooled data from 64 surveys of the original 71 surveys from the WHS due to slight differences in certain questions or high missing rates (>10%) for important covariates. The total number of respondents with responsiveness experiences (combining ambulatory and inpatient), who answered vignette questions for each set, was 150 632, comprised as follows: 38 331 for set A vignettes, 38 333 for set B vignettes, 37 953 for set C vignettes, and 36 015 for set D vignettes. Total responses are lower for set D because the Guatemala survey excluded set D vignettes for some unknown reason. Missing rates were similar across vignette sets. Average item missing rates of the dependent and independent covariates were low (less than 1%). Missing data were filled using the *mi* regression command in Stata (version 12), based on

2 covariates – country and inpatient or outpatient: averages of 10 imputations were rounded to the nearest ordinal response category for each individual.

| Vignette, rotation set A, dignity domain (vignette 1, best on scale: 1 to 5) |
|---|
| [Julia] was pregnant and went to the hospital coughing blood. A nurse welcomed her gently and helped her to a private room. A female doctor came to examine her and gave her a clean gown to replace her blood-stained clothes. |
| Vignette question |
| How would you rate his/her experience of being greeted and talked to respectfully? |

Figure 4.1 Exemplar of a dignity vignette and vignette question

Individual factors

In selecting individual-level factors for testing, we first considered which factors had been tested across the literature. Individual-level factors that had been tested for their impact on reporting heterogeneity in the responsiveness literature (formal and grey published between 2008 and 2012) have usually included: education or income, as markers of socioeconomic class or comprehension ability^{20,27,28}, and age, as a proxy of expectations related to intergenerational differences.^{10,29} Sex (gender) has been used to model gender differences in values and to model health reporting behaviour.³⁰ In the American literature, where the use of vignettes is less common in spite of the extensive patient experience research, patient-mix adjustment for comparing health plans adjusted for factors such as sex, health, ethnicity, age and education.¹²

Considering these and other studies, we finally adopted three categories of variables obtained from the World Health Survey: socio-demographic variables (category 1), health-related experience variables (category 2) and health value system variables (category 3). Category 1 includes: (1) education (ed) (0= primary or less, 1=secondary, 2=high school or more), (2) age (age) (18-30, 31-45, 45+), (3) sex (dichotomous, 1=male), (4) perceived health (continuous, 1 to 5); and one less commonly used variables: (5) alone (1, if the respondent was divorced, single or not cohabiting). Income was excluded as education provided a good substitute for a socioeconomic indicator and was more complete. Category 2 includes: (6) *intens* (1, having had an inpatient experience in the previous 5 years compared with only ambulatory experiences), (7) *careot* (1, caring for others with chronic illnesses). Category 3 includes: (8) *impresp* (1, own perceived importance of responsiveness as a health system goal).

Regression model

The analysis used mixed multilevel ordered probit models for (the ordinal) responses to each vignette question on each domain (5 vignettes on 8 domains, 40 regressions). The mixed nature of the model refers to the combination of a component consisting of a 'fixed' common set of variables, and a 'random' component. The multilevel model assumes, in this case, two levels with two different estimates of error variances— individual (1) and country (2). The level 1, 'fixed', effects refer to the coefficients on the individual-level covariates. As we operationalized age and education into two dichotomous variables, the original 8 variables were technically represented in the regressions by 10 covariates. At level 2, the country dummies were used to estimate the extent of error variance depending on the respondent's country. No further country-level covariates were included due to constraints in computational power and software, and the intended focus on exploring and reporting on individual-level factors in order to support within country comparability across population subgroups. The ordered probit model was used given that responsiveness vignettes were rated on a 1 to 5 ordinal scale from "very good" (1) to "very bad" (5).

In the ordered probit model, the scale cut-points were assumed fixed across respondents (no random error), only subject to modelling by individual factors (covariates). The model set the variance of the residuals at level 1 equal to 1, and estimated the residual (unexplained, random) variance at level 2 and the covariate coefficients (the 'betas', fixed effect). We further justified the use of a multilevel model by testing the log likelihood ratio for the model with and without individual-level covariates across several vignettes and domains, which generally rejected the null-hypothesis of $\beta=0$ at the $p=0.000$ level. We also used the log likelihood ratio test for a model using only category 1 variables, compared with that including category 2 and 3, which also rejected the null-hypothesis as above.

In two-level multilevel regression models, the residual variance at level 2 (in this analysis, a random intercept variance, $\text{var}(u)$), is a useful statistic for interpreting results. The relative size of the unexplained variance at level 1 and level 2 is an indicator of the importance of individual versus country factors, which relates to the final aim of this paper. The coefficient labelled 'rho', also known as the intraclass correlation coefficient or the variance partitioning coefficient, is an additional useful statistic that serves to partition the total unexplained variance. Rho answers the question of how much dependency there is in the data between individuals' reporting behaviour with respect to the survey they are in. Rho is calculated as the percentage of the total unexplained variance explained by $\text{var}(u)$, where total variance is estimated as a weighted sum of between variance ($\text{var}(u)$, for each case identifying the country average and taking the distance from mean) and individual variance ($\text{var}(e)$, individual case distances from the overall mean). Rho ranges between 0 and 1 but values of 0 are rarely seen.³¹

The size of ρ indicates how important it is to take country into account when judging the importance of individual-level predictors of differential reporting behaviour, or scale use, providing an overall assessment of the dependency of variation on country. Failure to use multilevel models when level 1 data is dependent on level 2 will bias the estimation of standard errors which in turn affects the significance of covariates. According to some sources, values considered high by convention in the social sciences are those around 0.3, while ρ s below 0.1 are considered not significant.³¹ In our analytical context 0.3 would imply that a third of the variance in vignette responses is systematically dependent on inter-country variation.

Effect patterns

The relative impact or effect of an individual-level covariate on the rating of a particular vignette can be interpreted by comparing covariate coefficients. A negative sign on a covariate coefficient implies that a particular group of respondents is more optimistic at a particular point on the scale (where the scale is ordered from 1, “very good”, to 5, “very bad”). Positive coefficients point to the reverse. When the coefficients for a particular covariate across all 5 vignettes of a particular domain are considered, their collective configuration can usually be described by a particular pattern as follows:

- I. *scale elongation*: present if respondents are more optimistic (negative sign on covariate coefficient) on good states in the domain (i.e. vignettes 1 and 2); and more pessimistic (positive sign on coefficient) on worse states (vignettes 4 and 5);
- II. *scale contraction*: where respondents are less optimistic on good states and less pessimistic on worse states; II is the inverse of I;
- III. *optimistic scale shift*: most coefficients (at least three out of five) have a negative sign and the remainder have coefficients close to zero (also known as parallel shifts or index shifts³²;
- IV. *pessimistic scale shift*: most coefficients (three or four out of five) have a positive sign; IV is the inverse of III.

Any other arrangement could be labelled, ‘mixed’ or ‘no pattern’. Effect strength can be distributed asymmetrically across the scale—stronger on a particular end of the scale (worse scenarios or better scenarios), or symmetrically—roughly the same size coefficients on the two best scenario and two worst scenario vignettes. There was no theory-based expectation on the predominance of symmetry or other patterns.^{10,33}

Reporting results

Tables 4.1, 4.2, and 4.3 summarize the main regression results for all 40 regressions. Table 4.1 shows the values for ρ . Tables 4.2 and 4.3 and Figure 4.2 display level 1 (individual-level) results and should be viewed jointly. Table 2 lists vignette regression covariate coefficients. Figure 4.2 illustrates the model results for one communication vignette (the best) by comparing respondents with different education levels). The predicted probabilities plot shows the relative size of both individual-level and country-level effects for a randomly drawn set of 9 countries.

Statistical software

The statistical software used for the regression analyses was the Generalized Linear Latent and Mixed Models (GLLAMM) software in Stata (version 12). The GLLAMM software enables the study of reporting behaviour using incrementally more complex multilevel models for categorical variables within a single statistical framework.^{34,35} All regressions were run initially with 6 quadrature integration points, then 8, then 12 and finally 14 integration points by saving the covariance matrix from each regression and using it in the subsequent regression. This integration procedure is recommended³⁵ to reduce processing time. Model estimates presented here were stable between 12 and 14 integration points.

RESULTS

Table 4.1 reports results on level 2 of the model for all 40 regressions: the intraclass correlation coefficient, referred to as ρ . The unexplained variation that is attributed to country-level factors ranges between 0.07/0.08 (see the fourth and second vignettes of social support, choice and autonomy domains) to 0.26 (vignette 5 (v5) of the quality of basic amenities domain). Relative to level 1, where unexplained variation is set to 1, country level unexplained variance is between one tenth and one third as large.

Across the 40 regressions, ρ ranges from a low of 0.07/0.08 (8 % of total residual variance) for choice vignette 3, to a high of 0.26 (26%) for the quality of basic amenities vignette 5. The size of ρ is an indicator of the how important it is to take country dependency into account when judging the importance of predictors (covariates) of scale behaviour changes at the individual-level. Country effects are more important for explaining heterogeneity in individuals' reporting behaviour in quality of basic amenities, prompt attention, and confidentiality domains, where three or more out of five vignette ρ 's are between 0.15 and 0.26. By this account, the effect of individual-level factors on reporting behaviour heterogeneity in dignity and social support domains are the least country dependent.

Table 4.1 Country-level rho results for all 40 vignette regressions

| Domain | v1 | v2 | v3 | v4 | v5 |
|----------------------------|------|------|------|------|------|
| Prompt attention | 0.16 | 0.20 | 0.09 | 0.09 | 0.18 |
| Dignity | 0.12 | 0.10 | 0.14 | 0.13 | 0.14 |
| Communication | 0.14 | 0.16 | 0.10 | 0.13 | 0.22 |
| Quality of basic amenities | 0.18 | 0.15 | 0.17 | 0.17 | 0.26 |
| Confidentiality | 0.16 | 0.12 | 0.15 | 0.16 | 0.23 |
| Choice | 0.17 | 0.12 | 0.08 | 0.12 | 0.15 |
| Social support (access to) | 0.14 | 0.08 | 0.12 | 0.12 | 0.13 |
| Autonomy | 0.14 | 0.08 | 0.12 | 0.07 | 0.16 |

Table 4.2 extracts the main results of the regressions and presents the 10 individual-level covariate coefficients for each set of 5 vignettes across 8 domains. Taking into account that the natural scale units for most covariates, except health, were 0 or 1, we observe that the effect size of covariates ranges from 0.00 (no effect) to 0.27 (substantial effect) for the individual-level covariate of high school completion or more years of education, for the communication vignette 5 (see line Com-v5). In order to examine the use of the responsiveness domain scale, it is necessary to assess the pattern of the covariate effects for all 5 vignettes. The pattern can be described in the first instance by the sign and size of the coefficients. For example, in considering a specific domain, autonomy, one observes that male respondents rate the best autonomy vignette (Aut-v1) on average 0.03 points higher. This means they are more pessimistic at the positive end of the scale, moving the response category for the best vignette from “very good” towards “good”, while rating the worst vignette (Aut-v5) on average -0.05 points lower: they are more optimistic at the bottom of the scale, moving the response category from “very bad” to “bad”. Therefore males use a more contracted set of scale values, relative to females, and their scale use pattern would be classified as pattern II. The largest coefficients are for education, caring for others with a chronic illness, own health and the importance of responsiveness. In view of the ordinal 1 to 5 scale units for self-reported health status, which was treated as a continuous variable, the strength of association of changes in reporting behaviour with health was on average equal to that for education (0,1 indicator variable). In some domain sets, where coefficients are shaded in grey, the range in coefficient size is smaller (0.00-0.03).

Table 4.2 Regression coefficients for individual covariates from 40 multilevel ordered probit regression models using the World Health Survey data: 10 covariates regressed on each vignette (response=1 “very good”, response=5 “very bad”) [using: bold and italics for $p \leq 0.05$, grey shading when coefficient size range for all vignettes in a domain is only | 0.01| to |0.03|]

| Domain and vignette (v) (v1=best vignette, v5 = worst vignette) | Category 1: Socio-demographic | | | | | Category 2: Health-related | | | Category 3: Health value system | |
|---|-------------------------------|----------------------------|--------------------------------------|------------------|--------------|------------------------------------|--|--|------------------------------------|--|
| | Male | Ed 1 (completed secondary) | Ed 2 (completed high school or more) | Age 1 (>30, <46) | Age 2 (>45) | Alone (marital/co-habiting status) | Caring for others with chronic illness | Intensity (inpatient visit last 5 yrs) | Health (1(very good)-5 (very bad)) | Responsiveness importance (more than health, financial protection) |
| PA-v1 | -0.01 | -0.06 | -0.12 | -0.00 | -0.00 | 0.01 | -0.07 | 0.02 | 0.05 | -0.04 |
| PA-v2 | 0.00 | 0.05 | 0.05 | -0.02 | -0.07 | 0.00 | -0.02 | 0.01 | 0.02 | 0.00 |
| PA-v3 | 0.00 | 0.05 | 0.05 | -0.02 | -0.07 | 0.00 | -0.02 | 0.01 | 0.02 | 0.00 |
| PA-v4 | 0.00 | 0.09 | 0.13 | -0.01 | -0.05 | 0.01 | 0.09 | -0.02 | -0.01 | 0.04 |
| PA-v5 | -0.01 | 0.08 | 0.15 | -0.01 | -0.02 | 0.01 | 0.10 | -0.03 | 0.00 | 0.05 |
| Dig-v1 | 0.01 | -0.08 | -0.16 | 0.02 | 0.02 | 0.00 | -0.08 | -0.01 | 0.03 | -0.05 |
| Dig-v2 | 0.00 | -0.05 | -0.11 | 0.01 | 0.03 | -0.01 | -0.03 | -0.02 | 0.03 | -0.04 |
| Dig-v3 | 0.00 | 0.03 | -0.01 | 0.01 | -0.02 | 0.01 | 0.00 | -0.05 | 0.01 | 0.00 |
| Dig-v4 | -0.04 | 0.07 | 0.10 | -0.01 | -0.04 | -0.01 | 0.04 | 0.01 | 0.01 | 0.04 |
| Dig-v5 | -0.02 | 0.08 | 0.18 | -0.01 | -0.06 | 0.01 | 0.06 | -0.02 | 0.00 | 0.04 |
| Com-v1 | 0.01 | -0.09 | -0.17 | -0.01 | 0.03 | 0.01 | -0.04 | 0.01 | 0.04 | -0.05 |
| Com-v2 | 0.01 | -0.09 | -0.11 | -0.03 | -0.02 | 0.00 | -0.03 | 0.03 | 0.04 | -0.03 |
| Com-v3 | -0.02 | 0.01 | 0.04 | 0.00 | -0.01 | -0.02 | 0.04 | -0.01 | 0.02 | 0.05 |
| Com-v4 | -0.04 | 0.07 | 0.16 | 0.01 | 0.03 | -0.03 | 0.04 | 0.00 | 0.01 | 0.04 |
| Com-v5 | -0.03 | 0.15 | 0.27 | 0.02 | 0.00 | -0.04 | 0.08 | 0.00 | 0.00 | 0.06 |
| QBA-v1 | 0.01 | -0.13 | -0.18 | -0.01 | 0.03 | 0.02 | -0.08 | 0.02 | 0.03 | -0.04 |
| QBA-v2 | 0.00 | -0.05 | -0.04 | -0.01 | 0.00 | 0.01 | 0.00 | -0.01 | 0.04 | -0.03 |
| QBA-v3 | -0.01 | 0.07 | 0.13 | 0.02 | -0.01 | -0.04 | 0.04 | -0.01 | 0.00 | 0.05 |
| QBA-v4 | -0.04 | 0.12 | 0.23 | 0.02 | 0.01 | -0.02 | 0.06 | -0.01 | 0.00 | 0.05 |
| QBA-v5 | -0.02 | 0.12 | 0.23 | 0.01 | -0.03 | -0.01 | 0.09 | -0.01 | 0.01 | 0.09 |
| Con-v1 | 0.02 | -0.10 | -0.17 | 0.01 | 0.03 | 0.00 | -0.03 | 0.00 | 0.02 | -0.07 |
| Con-v2 | -0.01 | 0.03 | 0.05 | 0.01 | -0.03 | 0.01 | 0.00 | -0.01 | 0.01 | 0.02 |
| Con-v3 | -0.04 | 0.06 | 0.10 | 0.03 | 0.00 | 0.01 | 0.01 | -0.01 | 0.00 | 0.05 |
| Con-v4 | -0.04 | 0.07 | 0.15 | 0.01 | -0.01 | -0.02 | 0.04 | -0.02 | -0.01 | 0.04 |
| Con-v5 | -0.03 | 0.04 | 0.18 | 0.02 | 0.01 | -0.01 | 0.05 | 0.02 | -0.01 | 0.05 |
| Ch-v1 | 0.02 | -0.09 | -0.20 | 0.01 | 0.00 | 0.02 | -0.01 | 0.00 | 0.03 | -0.05 |
| Ch-v2 | 0.01 | -0.02 | -0.07 | -0.02 | -0.02 | -0.02 | -0.04 | -0.01 | 0.03 | -0.05 |
| Ch-v3 | 0.00 | 0.03 | 0.06 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.02 |
| Ch-v4 | -0.01 | 0.10 | 0.19 | 0.02 | 0.01 | 0.00 | 0.01 | -0.01 | 0.01 | 0.02 |
| Ch-v5 | -0.01 | 0.03 | 0.11 | -0.01 | -0.05 | 0.00 | -0.02 | 0.00 | 0.01 | 0.05 |

Table 4.2 Regression coefficients for individual covariates from 40 multilevel ordered regression models using the World Health Survey data (continued)

| Domain and vignette (v) (v1=best vignette, v5 = worst vignette) | Category 1: Socio-demographic | | | | | Category 2: Health-related | | | Category 3: Health value system | |
|---|-------------------------------|----------------------------|--------------------------------------|------------------|--------------|------------------------------------|--|--|------------------------------------|--|
| | Male | Ed 1 (completed secondary) | Ed 2 (completed high school or more) | Age 1 (>30, <46) | Age 2 (>45) | Alone (marital/co-habiting status) | Caring for others with chronic illness | Intensity (inpatient visit last 5 yrs) | Health (1(very good)-5 (very bad)) | Responsiveness importance (more than health, financial protection) |
| SS-v1 | -0.01 | -0.05 | -0.12 | -0.01 | 0.03 | 0.00 | -0.01 | -0.02 | 0.04 | -0.04 |
| SS-v2 | -0.01 | -0.06 | -0.07 | -0.03 | -0.01 | 0.01 | -0.03 | 0.00 | 0.03 | -0.05 |
| SS-v3 | 0.02 | 0.00 | 0.02 | 0.00 | 0.02 | 0.00 | 0.00 | -0.01 | 0.02 | 0.01 |
| SS-v4 | -0.04 | 0.05 | 0.15 | 0.02 | -0.01 | -0.02 | 0.04 | -0.02 | 0.00 | 0.06 |
| SS-v5 | -0.03 | 0.05 | 0.11 | 0.03 | 0.00 | 0.00 | 0.06 | 0.00 | 0.01 | 0.04 |
| Aut-v1 | 0.03 | 0.01 | -0.06 | -0.01 | 0.02 | 0.03 | -0.03 | -0.02 | 0.03 | -0.05 |
| Aut-v2 | 0.02 | -0.05 | -0.04 | -0.02 | -0.05 | 0.00 | -0.01 | -0.01 | 0.04 | -0.04 |
| Aut-v3 | 0.00 | 0.05 | 0.05 | 0.01 | 0.01 | -0.02 | 0.01 | 0.01 | 0.01 | 0.00 |
| Aut-v4 | -0.02 | -0.03 | 0.00 | -0.02 | -0.05 | 0.00 | 0.03 | 0.00 | 0.02 | 0.02 |
| Aut-v5 | -0.05 | 0.07 | 0.13 | 0.01 | -0.03 | -0.01 | 0.05 | -0.03 | 0.00 | 0.06 |

Key: PA: prompt attention domain; Dig: dignity domain; Com: communication domain; QBA: quality of basic amenities domain; Con: confidentiality domain; Ch: choice domain; SS: social (family and community) support domain; Aut: autonomy domain

Table 4.3 categorizes the scale use patterns observed for each covariate and domain combination for each set of vignettes as: scale elongations (I), scale contractions (II), or positive (optimistic) (III) or negative (pessimistic) shifts (IV). Starting with the first covariate listed in column 1, we observe that reporting scales used by males are generally more contracted. More education is associated with scale elongation (I) across all domains. Age patterns are less consistent across domains but there is a slight tendency for older respondents to display an optimistic scale shift (typology III). For marital status (alone), we observe scale contraction (II), and a negative index shift (IV), but most effect sizes are close to zero. The coefficients of intensity (1, having had an inpatient visit in the previous 5 years) exhibit scale contraction (II) and optimistic shifts (III), but again effect sizes are close to zero. Caring for family/friends with chronic illness/disability is associated with scale elongation (I) in all domains. Poorer own health status (higher value) is associated with a pessimistic scale index shifts (IV) in all domains except confidentiality. If responsiveness is considered more of a priority (most important), the reporting scale is elongated (I). The most common finding on symmetry is for effects to be asymmetric with stronger effects found on the worse vignettes.

Table 4.3 World Health Survey reporting behaviour heterogeneity patterns: classified from the analysis of covariate coefficients for each domain (see Table 4.2) [using: grey shading when coefficient size range for all vignettes in a domain is only | 0.01| to |0.03|]

| Domain and vignette (v) (v1=best vignette, v5 = worst vignette) | Category 1: Socio-demographic | | | | | | Category 2: Health-related | | | Category 3: Health value system |
|---|-------------------------------|----------------------------|--------------------------------------|------------------|-------------|------------------------------------|--|---|------------------------------------|--|
| | Male | Ed 1 (completed secondary) | Ed 2 (completed high school or more) | Age 1 (>30, <46) | Age 2 (>45) | Alone (marital/co-habiting status) | Caring for others with chronic illness | Intensity (in-patient visit last 5 yrs) | Health (1(very good)-5 (very bad)) | Responsiveness importance (more than health, financial protection) |
| Prompt attention | No pattern | I | I | III | III | IV | I | II | IV | I |
| Dignity | II | I | I | II | II | IV | I | III | IV | I |
| Communication | II | I | I | I | No pattern | II | I | II | IV | I |
| Quality of basic amenities | II | I | I | No pattern | II | II | I | II | IV | I |
| Confidentiality | II | I | I | IV | No pattern | II | I | No pattern | II | I |
| Choice | II | I | I | III | No pattern | IV | III | No pattern | IV | I |
| Social support | III | I | I | I | III | II | I | III | IV | I |
| Autonomy | II | No pattern | I | III | III | II | I | III | IV | I |

Key: I. scale elongation; II. scale contraction; III. optimistic scale shift (negative signs for coefficients); IV. pessimistic scale shift (positive signs for coefficients)

Figure 4.2 plots the effect of education on the rating scale for positive scenarios of communication, in nine countries (Estonia, Ethiopia, Finland, France, England, Georgia, Ghana, Greece, and Guatemala). The probability of assigning a lower rating than “very good” to the best communication vignette (com-v1) is compared within each country, for respondents of lower education (primary or less, 0) with those with high school or more years of education. From Figure 4.2, taking France (FRA) as an example, we observe that predicted probabilities of scoring com-v1 (communication, vignette 1 (best)) as less than “very good” (i.e. “good”, “moderate”, “bad” or “very bad”) are higher in lower educated groups - ranging from 0.40 in 0.55, and lower in more

educated groups – ranging from 0.35 to 0.45 (category 1 on the x-axis). These ranges also vary across countries.

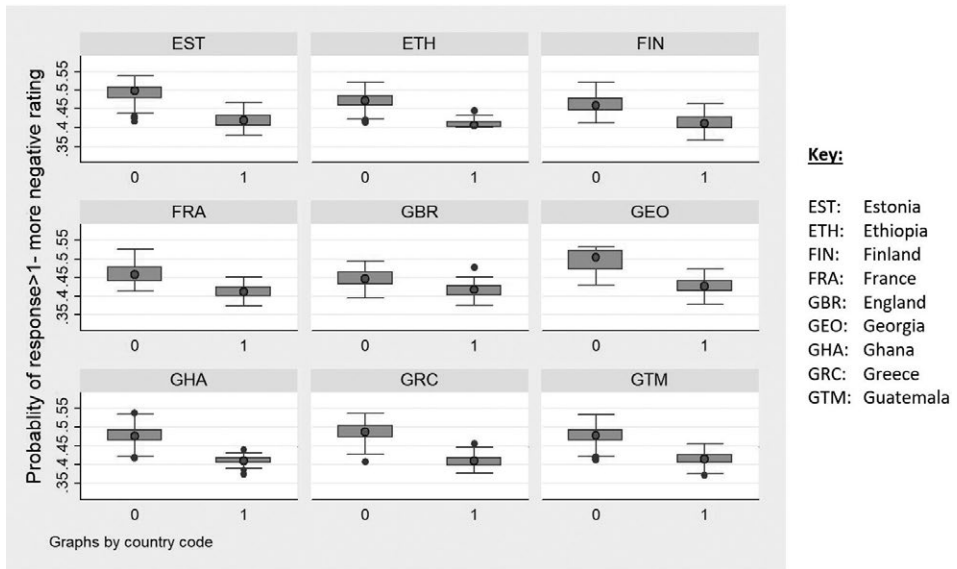


Figure 4.2 Predicted probabilities from regression models for the communication domain: the probability of an individual, according to education (0, <=primary; 1,>primary), rating the best communication vignette (com-v1) more ‘pessimistically’ or ‘negatively’ for a random sample of 9 countries

DISCUSSION

We succeeded in applying a comprehensive strategy to describe the effect of 8 individual-level factors on reporting behaviour heterogeneity for responsiveness, using cross-sectional data from 64 countries. The two-level regression approach revealed higher unexplained variation at level 1 (individual) than at level 2 (country). The effect of individual-level covariates on reporting behaviour is strongly independent of country in 2 out of 8 domains (dignity and social support) and strongly dependent on country in 3 out of 8 domains of prompt attention, quality of basic amenities and confidentiality). Effect patterns differ across individual-level factors but are consistent for a particular factor across domains. Factors displaying elongation patterns have coefficients with larger effect sizes than those with shift patterns. Effect sizes are generally asym-

metric, with vignettes describing worse states having more variance explained by the factors that were tested.

The factors tested grouped individual-level variables into three categories. In category 1 (socio-demographic), education has the strongest effects on reporting behaviour heterogeneity; in category 2 (health-related experiences), perceived health state, followed by caring for others with a chronic illness, have the strongest effects; and in category 3 (health value system), the newly introduced variable of importance of responsiveness has a strong effect. Commonly used variables, namely, age and sex, have smaller effects than these newly tested variables.

Our paper elaborates on Lindeboom and Doorslaer³² in explicitly describing effect patterns. The effects of covariates on reporting behaviour heterogeneity are described in terms of four basic patterns, all of which were observed: (I) contraction; (II) elongation; (III) optimistic shift; and (IV) pessimistic shift. In addition effects are described in terms of strength, consistency across domains, and symmetry. Education, caring for others with a chronic illness, and the importance of responsiveness, displayed patterns I. Perceived health (state) displayed pattern IV. While a complete comparison with other literature is not possible, as no other study has covered as many variables, domains and countries, comparisons are possible for particular domains and for particular covariates–domain combinations, as provided below.

Overall, our findings for category 1 variables that have been covered in other studies are similar. The effect of education on reporting behaviour is the most well established effect in the literature as referred to earlier, and is confirmed by our large study. Other studies also found the education effect characterised by scale elongation. In addition, we found asymmetric stronger effects in the negative scenarios.

Sirven et al.¹⁰ used the SHARE dataset to test individual-level effects for sex, age, and health variables. Their data were from elderly people (50 years and over) from 11 countries (Germany, Sweden, Netherlands, Spain, Italy, France, Denmark, Greece, Belgium, Czech Republic, Poland) in Europe, and included 3 domains. The domains: time to wait for medical treatment, conditions of the health facilities and communication with doctors, are roughly equivalent to prompt attention, quality of basic amenities, and communication domains in our study. With regard to sex, their study found covariate–domain specific patterns. There was no effect in prompt attention; a slight effect in the quality of basic amenities domain (males less critical of more positive scenarios), and in the communication domain (males more critical of worse scenarios). In our results, covering more covariates and domains, covariate-specific patterns for different domains were less important relative to the importance of the general patterns that emerged for covariates across all domains. With regard to general patterns for age, Sirven et al.¹⁰ describe only a slight

scale shift effect in communication and prompt attention domains (coefficients range 0.003-0.007, $p < 0.05$). Our data contained a wider age range of respondents and showed weak effects for the age group of 30-45 years, but stronger effects for the over 45 years age group.

With regard to health states, Sirven et al.¹⁰ investigated specific health conditions, whereas our study focused on general perceived health (status). They found that depressed respondents exhibited pessimistic scale shifts in the 3 domains, particularly for the vignettes describing worse scenarios. The overall health state covariate in our study also produced a pessimistic shift across 8 domains but stronger effects were present in better scenarios for health. (Although more generally across other covariates, stronger effects were present on worse scenarios). For use of inpatient versus ambulatory services, our results showed small effects as in Sirven et al.¹⁰

Our study has several limitations. First, the dual extension of both individual-level factors and modelling country level effects was not possible due to a lack of computational power and software limitations. We opted to focus on the extension of individual-level factors, while maintaining an effect for countries, in order to concentrate on within country comparability challenges. Consequently, no specific country-level covariates were included in our analysis. They warrant consideration in future analyses simultaneously with an extended set of individual-level variables if possible. Other studies that have focused on country-level factors have examined civil liberties, health expenditure per capita and culture.⁹ But in view of the fairly large unexplained variance still found at the individual-level, extending the individual-level covariates is also warranted e.g., by limiting countries and expanding the number of covariates. At the individual-level we could then add, from categories 1, 2 and 3 respectively: employment status (not working for pay, or informal versus formal), different domains of health (affect, pain, functioning), and satisfaction with how health care runs. We expect employment status to affect expectations, while health affect can change moods, fixing attitudes and thereby affect reporting behaviour. We expect satisfied people to be more likely to rate positive scenarios optimistically (the 'halo effect').

Second, our model assumes no correlation between the individual-level factors and the country-level random effect, while in fact there is likely to be some correlation.³⁶ In particular, the variable, own perceived importance of responsiveness as a health system goal, may be influenced by culture. Further comparisons of models could be developed to test this further. Our preliminary assessment is that it may only be relevant for some types of countries, thereby limiting any effect on the estimates presented in this paper.

Third, inter-country heterogeneity in inter-rater reliability and the different sizes of countries could have influenced our findings. For the latter, as only 1 country, Mexico, had a much larger size than the others, it likely that the diversity of the other countries would have balanced out any

singular impact this country had on the results. Fourth, our study uses quantitative data only. Further exposition of the mechanisms underpinning reporting behaviour of actual experiences, alluded to in the introduction, require further qualitative, longitudinal and experimental data.

The evidence has important implications for the use of vignettes in surveys e.g., the minimum number (a cost issue), and their optimal placement on the latent scale. Scale patterns are less in favour of scale shifts and more in favour of scale elongation or contraction; most effects are asymmetric; and most are similar across domains. We would therefore consider that most information on reporting behaviour heterogeneity could be extracted from using just 3 vignettes—one on the more favourable scenarios (“good”) and 2 on the worse scenarios (“bad” and “very bad”). As covariate effects are consistent across domains, a further cost-saving measure for survey design is to combine several domains at a point in the scale into single vignettes, but further research would be needed to assess this option.

A frequently asked policy question within countries is: “how well are services performing for more advantaged relative to less advantaged individuals?” Disclosing reporting behaviour heterogeneity is particularly important when the variables used to identify less advantaged individuals and populations also shape the use of the response scale. Prior to this study, how a wide range of factors affected reporting behaviour heterogeneity in response to responsiveness survey questions, was not known with as much certainty. Using knowledge of these patterns gleaned from 64 countries can improve comparability of health services or plans within countries and improve the confidence of national policy-makers in user-assessed health quality metrics. Given the role played by individual-level factors, paying attention to these when reporting on responsiveness, will enhance both within and across-country comparability. Regardless, a reporting convention that explicitly describes the main sources of ‘bias’ alongside any ‘adjusted’ performance values seems to be the right balance to strike, regardless of adjustment technique or setting.

REFERENCES

1. Anand S, Ammar W, Evans T, Hasegawa T, Kissimova-Skarbek K, Langer A, et al. Report of the scientific peer review group on health systems performance assessment. In: Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism*. Geneva: World Health Organization; 2003. p. 839–916.
2. Evans D, Hsu J, Boerma T. Universal health coverage and universal access. *B World Health Organ*. 2013;91:546–A.
3. Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bull World Health Organ*. 2000;78(6):717–31.
4. Gostin L, Hodge J, Valentine N, Nygren-Krug H. The domains of health responsiveness: a human rights analysis. *Health and human rights working paper no. 2*. Geneva: World Health Organization; 2003. [online: http://apps.who.int/iris/bitstream/10665/73926/1/HHRWPS2_eng.pdf]
5. Anhang Price R, Elliott MN, Zaslavsky AM, Hays RD, Lehrman WG, Rybowski L, et al. Examining the role of patient experience surveys in measuring health care quality. *Medical Care Res Rev*. 2014;71(5):522–54.
6. Shiferaw S, Spigt M, Godefrooi M, Melkamu Y, Tekie M. Why do women prefer home births in Ethiopia? *BMC Preg Childbirth*. 2013;13:5.
7. Zimlichman EI, Rozenblum R, Millenson ML. The road to patient experience of care measurement: lessons from the United States. *Isr J Health Pol Res*. 2013;17(2):1.
8. Murray CJL, Kawabata K, Valentine NB. People's experience versus people's expectations. *Health Affair*. 2001;20(3):41–3.
9. Rice N, Robone S, Smith P. Vignettes and health systems responsiveness in cross-country comparative analyses. *J R Stat Soc Ser A*. 2012;175(2):337–69.
10. Sirven N, Santos-Eggimann B, Spagnoli J. Comparability of health care responsiveness in Europe. *Soc Indic Res*. 2012;105:255–71.
11. Kok R, Avendano Pabon M, Bago d'Uva T, Mackenbach J. Can reporting heterogeneity explain differences in depressive symptoms across Europe? *Soc Indic Res*. 2012;105:187–90.
12. Hargraves JL, Wilson IB, Zaslavsky A, James C, Walker JD, Rogers G, et al. Adjusting for patient characteristics when analyzing reports from patients about hospital care. *Med Care*. 2001;39(6):635–41.
13. de Silva A. A framework for measuring responsiveness. *Global programme on evidence for health policy discussion paper no. 32*. Geneva: World Health Organization; 2000. Accessed on 2 April 2016 from <http://www.who.int/healthinfo/paper32.pdf>.
14. Farley DO, Elliott MN, Haviland AM, Slaughter ME, Heller A. Understanding variations in Medicare consumer assessment of health care providers and systems scores: California as an example. *Health Serv Res*. 2011;46(5):1646–62.
15. Groves R. *Survey error models and cognitive theories of response behaviour*. Cognition and survey research Wiley series in probability and statistics. New York: Wiley; 1999. p. 235–50.
16. Jasso G. Factorial survey methods for studying beliefs and judgments. *Sociol Method Res*. 2006 February;34(3):334–423.

Chapter 4. Health System Responsiveness by N Valentine

17. Blendon R, Schoen C, DesRoches C, Osborn R, Zaper K. Common concerns amid diverse systems: health care experiences in five countries. *Health Affair*. 2003;22(3):106–21.
18. Holzer BM, Minder CE. A simple approach to fairer hospital benchmarking using patient experience data. *Int J Qual Health Care*. 2011;23(5):524–30.
19. Cox DR, 2011 (August 23). Discussion on the paper by Rice, Robone and Smith (2012). In VS Isham (Chair), Vignettes and health systems responsiveness in cross-country comparative analyses. The 58th World Royal Statistical Society Annual Meeting, Dublin, Ireland.
20. Rice N, Robone S, Smith P. Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness. *Eur J Health Econ*. 2011;12:141–62.
21. King N, Harper S, Young M. Who cares about health inequalities? Cross-country evidence from the World Health Survey. *Health Policy Plann*. 2013;28(5):558–71.
22. Szwarcwald C, Souza-Júnior P, Damacena G. Socioeconomic inequalities in the use of outpatient services in Brazil according to health care need: evidence from the World Health Survey. *BMC Health Serv Res*. 2010;10(1):1–7.
23. Valentine NB, Prasad A, Rice N, Ribone S, Chatterji S. Health systems responsiveness: a measure of the acceptability of health-care processes and systems from the user's perspective. In: Smith P, Mossialos E, Papanicolas I, Leatherman S, editors. *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press; 2010.
24. King G, Murray CJL, Salomon J, Tandon A. Enhancing the validity and cross-cultural comparability of measurement in survey research. *Am Polit Sci Rev*. 2004 (February);98(1):191–207.
25. Grol-Prokopczyk H, Freese J, Hauser RM. Using anchoring vignettes to assess group differences in general self-rated health. *J Health Soc Behav*. 2011;52(2):246–61.
26. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet*. 2007;370:851–8.
27. Rice N, Robone S, Smith P. Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness. *Eur J Health Econ*. 2011;12:141–62.
28. Malhotra C, Kyung Do Y. Socio-economic disparities in health system responsiveness in India. *Health Policy and Planning*. 2013;28(2):197–205.
29. Kowal P, Naidoo N, Williams S, Chatterji S. Performance of the health system in China and Asia as measured by responsiveness. *Health*. 2011;3(10):638–45.
30. Peracchi F, Rossetti C. Heterogeneity in health responses and anchoring vignettes. *Empir Econ*. 2012;42:513–38.
31. Pillinger R. An advanced user's guide to Stat-JR version 1.0.0. Measuring dependency Bristol: University of Bristol; 2013 [cited 2014 February 5]. A transcript of measuring dependency presentation. [online: <http://www.bristol.ac.uk/cmm/learning/videos/measuring-dep.html#top>]
32. Lindeboom M, van Doorslaer E. Cut-point shift and index shift in self-reported health. *J Health Econ*. 2004;23:1083–99.
33. Jürges H, van Soest A. Comparing the well-being of older Europeans: introduction. *Soc Indic Res*. 2012, May;105:187–90.

34. Rabe-Hesketh S, Skrondal A, Pickles A. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *J Econometrics*. 2005;128:301–23.
35. Rabe-Hesketh S, Skrondal A. *Multilevel and longitudinal modeling using Stata*. College Station: Stata Press; 2012.
36. Clark TS, Linzer DA. Should I use fixed or random effects? *The society for political methodology. Political Science Research and Methods*. 2015;3(2):399–408.

Appendix 4.1 Exemplars of vignettes for the domain of dignity, and vignette questions for all domains, from the World Health Survey

Five vignettes from rotation set A, dignity domain, ranked from best to worst

[Julia] was pregnant and went to the hospital coughing blood. A nurse welcomed her gently and helped her to a private room. A female doctor came to examine her and gave her a clean gown to replace her blood-stained clothes.

[Conrad] had bad flu. He went to the clinic. The nurse expressed concern about [Conrad]'s cough and called the doctor, who gave [Conrad] a full chest examination behind a large screen that hid him from the view of other patients.

[Patricia] went to a crowded clinic. At first, no-one greeted her but after waiting for 5 minutes a nurse called her to the examination area where she was examined behind a small screen that mostly hid her from the other patients.

[Anya] took her baby for a vaccination. The nurse said hello and but did not ask for [Anya's] or the baby's name. The nurse also examined [Anya] and made her remove her shirt in the waiting room.

[Said] has AIDS. When he goes to his health centre the nurses do not talk to him and deliberately ignore him. During examinations, his clothes are removed and he is made to wait, half-naked in the waiting room.

| Domain | Vignette question wording for Set A, B, C, D |
|-------------------------------|--|
| A: Prompt attention | How would you rate the amount of time s/he waited before being attended to? |
| A: Dignity | How would you rate his/her experience of being greeted and talked to respectfully? |
| B: Communication | How would you rate his/her experience of how clearly health care providers explained things to him/her? |
| B: Quality of Basic Amenities | How would you rate the cleanliness of the rooms inside the facilities, including toilets? |
| C: Confidentiality | How would you rate the way the health services ensure s/he could talk privately to health care providers? |
| C: Choice | How would you rate his/her freedom to choose his/her health care provider |
| D: Social Support | For his/her last hospital stay, how would you rate the ease of having friends or family visit him/her? |
| D: Autonomy | How would you rate his/her experience about being involved in making decisions about his/her health care or treatment? |

Appendix 4.2 Reporting behaviour heterogeneity in prompt attention vignette 1 (best scenario): percentage of respondents answering “very good” stratified by (i) education (category 1); (ii) caring for others with chronic illness (category 2); (iii) judgement of responsiveness as most important (category 3) ^a

| Country ^b | Education (completed high school or more) (%) | | Caring for others with chronic illness (%) | | Responsiveness most important (vs. health outcome or financial protection) (%) | |
|------------------------|---|------|--|------|--|------|
| | No | Yes | No | Yes | No | Yes |
| Average | 0.54 | 0.58 | 0.54 | 0.56 | 0.54 | 0.56 |
| United Arab Emirates | 0.60 | 0.66 | 0.63 | 0.67 | 0.65 | 0.59 |
| Austria | 0.78 | 0.87 | 0.78 | 0.86 | 0.77 | 0.84 |
| Burkina Faso | 0.36 | 0.48 | 0.35 | 0.43 | 0.36 | 0.40 |
| Bangladesh | 0.42 | 0.54 | 0.43 | 0.40 | 0.42 | 0.42 |
| Bosnia and Herzegovina | 0.58 | 0.50 | 0.57 | 0.63 | 0.58 | 0.50 |
| China | 0.49 | 0.58 | 0.49 | 0.59 | 0.52 | 0.44 |
| Cote d'Ivoire | 0.31 | 0.32 | 0.33 | 0.21 | 0.29 | 0.38 |
| Congo | 0.41 | 0.44 | 0.45 | 0.31 | 0.42 | 0.38 |
| Comoros | 0.24 | 0.36 | 0.29 | 0.09 | 0.20 | 0.39 |
| Czech Rep. | 0.81 | 0.87 | 0.81 | 0.93 | 0.86 | 0.77 |
| Germany | 0.75 | 0.81 | 0.76 | 0.82 | 0.78 | 0.72 |
| Denmark | 0.70 | 0.84 | 0.77 | 0.76 | 0.76 | 0.78 |
| Dominican Republic | 0.29 | 0.28 | 0.30 | 0.26 | 0.30 | 0.27 |
| Ecuador | 0.52 | 0.55 | 0.53 | 0.50 | 0.52 | 0.53 |
| Spain | 0.59 | 0.65 | 0.59 | 0.65 | 0.59 | 0.63 |
| Estonia | 0.73 | 0.80 | 0.80 | 0.74 | 0.72 | 0.85 |
| Ethiopia | 0.66 | 0.90 | 0.64 | 0.84 | 0.66 | 0.58 |
| Finland | 0.52 | 0.68 | 0.58 | 0.65 | 0.62 | 0.60 |
| France | 0.72 | 0.73 | 0.75 | 0.63 | 0.72 | 0.74 |
| UK (England) | 0.74 | 0.79 | 0.73 | 0.83 | 0.71 | 0.82 |
| Georgia | 0.47 | 0.54 | 0.55 | 0.44 | 0.49 | 0.64 |
| Ghana | 0.71 | 0.57 | 0.71 | 0.68 | 0.70 | 0.73 |
| Greece | 0.70 | 0.63 | 0.67 | 0.65 | 0.67 | 0.66 |
| Guatemala | 0.25 | 0.36 | 0.26 | 0.27 | 0.26 | 0.27 |
| Croatia | 0.80 | 0.81 | 0.82 | 0.74 | 0.79 | 0.83 |
| India | 0.27 | 0.34 | 0.26 | 0.38 | 0.29 | 0.27 |
| Ireland | 0.57 | 0.72 | 0.56 | 0.84 | 0.58 | 0.62 |
| Israel | 0.69 | 0.72 | 0.70 | 0.74 | 0.73 | 0.70 |
| Italy | 0.47 | 0.55 | 0.49 | 0.55 | 0.48 | 0.55 |
| Kazakhstan | 0.48 | 0.47 | 0.44 | 0.58 | 0.47 | 0.47 |
| Kenya | 0.76 | 0.75 | 0.77 | 0.73 | 0.75 | 0.77 |

Appendix 4.2 continued

| Country ^b | Education (completed high school or more) (%) | | Caring for others with chronic illness (%) | | Responsiveness most important (vs. health outcome or financial protection) (%) | |
|----------------------|---|------|--|------|--|------|
| | No | Yes | No | Yes | No | Yes |
| Laos | 0.36 | 0.40 | 0.37 | 0.36 | 0.37 | 0.36 |
| Sri Lanka | 0.47 | 0.52 | 0.48 | 0.49 | 0.49 | 0.47 |
| Luxemburg | 0.46 | 0.67 | 0.52 | 0.57 | 0.52 | 0.58 |
| Latvia | 0.76 | 0.79 | 0.78 | 0.72 | 0.82 | 0.65 |
| Morocco | 0.87 | 0.83 | 0.86 | 0.87 | 0.84 | 0.88 |
| Mexico | 0.36 | 0.39 | 0.36 | 0.38 | 0.37 | 0.35 |
| Mali | 0.49 | 0.50 | 0.51 | 0.45 | 0.48 | 0.67 |
| Myanmar | 0.32 | 0.39 | 0.30 | 0.40 | 0.31 | 0.38 |
| Mauritania | 0.22 | 0.29 | 0.21 | 0.33 | 0.22 | 0.30 |
| Mauritius | 0.53 | 0.69 | 0.56 | 0.53 | 0.52 | 0.59 |
| Malawi | 0.66 | 0.58 | 0.65 | 0.71 | 0.63 | 0.71 |
| Malaysia | 0.34 | 0.41 | 0.37 | 0.38 | 0.38 | 0.35 |
| Namibia | 0.31 | 0.50 | 0.29 | 0.44 | 0.31 | 0.34 |
| Netherlands (The) | 0.65 | 0.54 | 0.57 | 0.51 | 0.57 | 0.52 |
| Norway | 0.77 | 0.81 | 0.75 | 0.85 | 0.78 | 0.81 |
| Nepal | 0.27 | 0.27 | 0.24 | 0.34 | 0.26 | 0.28 |
| Pakistan | 0.41 | 0.40 | 0.41 | 0.41 | 0.42 | 0.36 |
| Philippines | 0.24 | 0.23 | 0.25 | 0.20 | 0.21 | 0.30 |
| Portugal | 0.41 | 0.62 | 0.47 | 0.33 | 0.50 | 0.35 |
| Paraguay | 0.49 | 0.61 | 0.49 | 0.54 | 0.53 | 0.48 |
| Russia | 0.65 | 0.75 | 0.70 | 0.74 | 0.69 | 0.77 |
| Senegal | 0.45 | 0.52 | 0.45 | 0.51 | 0.44 | 0.55 |
| Slovakia | 0.72 | 0.85 | 0.81 | 0.82 | 0.82 | 0.78 |
| Slovenia | 0.83 | 0.76 | 0.80 | 0.78 | 0.81 | 0.77 |
| Sweden | 0.79 | 0.78 | 0.75 | 0.84 | 0.77 | 0.82 |
| Swaziland | 0.50 | 0.38 | 0.48 | 0.47 | 0.46 | 0.55 |
| Trinidad and Tobago | 0.33 | 0.29 | 0.33 | 0.32 | 0.29 | 0.53 |
| Tunisia | 0.61 | 0.59 | 0.58 | 0.66 | 0.61 | 0.58 |
| Ukraine | 0.52 | 0.62 | 0.61 | 0.60 | 0.59 | 0.64 |
| Uruguay | 0.57 | 0.59 | 0.57 | 0.59 | 0.58 | 0.55 |
| Viet Nam | 0.44 | 0.49 | 0.46 | 0.43 | 0.45 | 0.45 |
| South Africa | 0.48 | 0.44 | 0.48 | 0.42 | 0.48 | 0.43 |
| Zambia | 0.63 | 0.57 | 0.61 | 0.69 | 0.59 | 0.69 |

^a Tables on the other covariates and for the other 39 vignettes, are available from the authors on request; ^b Ordering is alphabetical, according to WHO standard abbreviation for country name

Appendix 4.3 World Health Survey vignette results: typology of reporting behaviour with respect to effect patterns and symmetry of effect strength, by covariate and domain

| Covariate scale use patterns | Stronger effect on worse scenarios | Stronger effect on better scenarios | Symmetric strength of effect | Stronger effect middle scenarios | No pattern |
|------------------------------|---|---|--|----------------------------------|---|
| I-elongation | Ed1-Com; Ed2-PA; Ed2-Dig; Ed2-Com; Ed2-QBA; Ed2-Con; Ed2-SS; Careot-PA; Careot-Dig; Careot-Com; Careot-QBA; Careot-Con; Careot-SS; Careot-Aut | | Ed1-PA; Ed1-Dig; Ed1-QBA; Ed1-SS; Ed1-Ch; Ed2-Ch; Ed2-Aut; Impresp-PA; Impresp-Dig; Impresp-Com; Impresp-QBA; Impresp-Con; Impresp-Ch; Impresp-SS; Impresp-Aut | | |
| II-contraction | Male-Dig; Male-Com; Male-QBA; Male-Con; Age2-Dig; Alone-Com; Alone-QBA; | | Male-Ch; Male-Aut; Age2-QBA; Health-Con | | |
| III-optimistic shift | Male-SS | Careot-Ch | Age2-PA; Age2-SS; Age2-Aut | Intens-Dig | |
| IV-pessimistic shift | | Health-PA; Health-Dig; Health-Com; Health-QBA; Health-Ch; Health-SS; Health-Aut | | | |
| No pattern | | | | | Male-Pa; Ed1-Aut; Age2-Com; Age2-Con; Age2-Ch |

Key:

Covariates

Male: male; Age1: >30, <46; Age2: >45; Alone: marital, co-habiting status; Careot: caring for others with chronic illness; Ed1: completed secondary schooling; Ed2: completed high school or more; Health: self-rated 1(very good)-5 (very bad); Impresp: responsiveness importance (more than health, financial protection)

Domains

Aut: autonomy; Ch: choice; Com: communication; Con: confidentiality; Dig: dignity; PA: prompt attention; QBA: quality of basic amenities; SS: social support

The background consists of several overlapping, semi-transparent grey planes that create a 3D effect. These planes are oriented in various directions, some parallel to the ground and others at an angle. Scattered across these planes are various text labels in a light grey, sans-serif font. The labels include: 'DIGNITY' at the top; 'PROMPT ATTENTION' on a plane below it; 'COMMUNICATION' on a plane further down; 'QUALITY BASIC A' and 'NITI' on a plane below that; 'CONFIDENTIALITY' on a plane below that; 'ACCESS TO SUPPORT NETWORKS' on a plane below that; 'AUTONOMY' on a plane below that; and 'CHOICE OF CARE PROVIDER' on the bottom-most plane. The overall composition is clean and modern, with a focus on geometric shapes and text.

PART II

Explaining why responsiveness matters
for people, services and policy

CHAPTER 5

Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of "health systems responsiveness" in 41 countries



N.B. Valentine, C. Darby, G.J. Bonsel
Social Science & Medicine. 2008; 66:1939-50

ABSTRACT

Quality of care research has reached some agreement on concepts like structure, process and outcome, and non-clinical versus clinical processes of care. These concepts are commonly explored through surveys measuring patient experiences, yet few surveys have focused on patient, or “user”, priorities across different quality dimensions. Population surveys on priorities can contribute to, although not replace participation in, policy decision making. Using 105,806 survey interview records from the World Health Organization’s (WHO’s) general population surveys in 41 countries, this paper describes the relative importance of 8 domains in the non-clinical quality of care concept WHO calls “health systems responsiveness”. Responsiveness domains are divided into interpersonal domains (dignity, autonomy, communication and confidentiality) and structural domains (quality of basic amenities, choice, access to social support networks and prompt attention). This paper explores variations in domain importance by country level variables (country of residence, human development, health system expenditure, and ‘geographic zones’) and by subpopulations defined by sex, age, education, health status, and utilization. Most respondents selected prompt attention as the most important domain. Dignity was selected second, followed by communication. Access to social support networks was identified as the least important domain. In general, convergence in rankings was stronger across subpopulations within countries than across countries. Yet even across diverse countries, there was more convergence than divergence in views. These results provide a ranking of quality of care criteria for consideration during health reform processes further to the usual emphasis on clinical quality and supply side efficiency.

INTRODUCTION

The quality of care literature supports the view that non-clinical aspects of health care are important dimensions of service quality.¹⁻⁴ In spite of interest in users' experiences of quality of care, Sitzia and Wood⁴ observed that "little, empirical data on the importance of the(se) components relative to each other has been published" (page 1838). Apart from being of theoretical interest, the relative importance of different components of quality of care communicates meaningful information. The knowledge can contribute to social dialogue about what is important in health services^{5,6}, it can be used to construct composite indices measuring users' experiences with health services⁷; and it may help to anticipate the popularity of health system changes. Health policy changes inevitably favour one or more aspect of non-clinical quality of care over another. It follows, therefore, that differences between population groups within a country can lead to conflict, and across countries to misinterpretation of inter-country comparisons made with composite measures.

The World Health Organization's (WHO's) 2000 World Health Report, expressed a serious concern with exploring users' priorities with respect to different aspects of health services, including for a non-clinical quality of care concept called "health systems responsiveness".⁸ The responsiveness concept covered 7 'domains', to which an eighth was subsequently added.⁹⁻¹¹ The 8 domains were categorized as either "respect for persons" domains (covering dignity, autonomy, confidentiality, and (clear) communication)¹, and "client orientation" domains (covering choice of care provider, prompt attention, quality of basic amenities, access to social support networks (during inpatient care)²). WHO stated that these non-clinical domains were of 'universal' importance, meaning important to all human beings, regardless of culture, sex, age and so on, a highly controversial claim at the time.¹²

Following this conceptual work, WHO launched the measurement of people's experiences with health services (labelled 'performance'), and people's views on the relative importance of these characteristics within a health system (labelled 'importance'), through the Multi-Country Survey Study on Health and Health Systems Responsiveness (the 'MCS Study'). MCS questionnaires were administered to households in 70 general population surveys in 60 countries.¹³ In 41 countries, questionnaires were administered by interviewers. In the remaining cases, surveys were self-administered (either delivered by hand, or postal). This paper focuses on the interviewer administered surveys (39 face-to-face and 2 telephone surveys (Canada, Luxembourg)), to reduce differences in bias associated with administration mode.^{14,15}

Previous studies have proclaimed there to be divergences in priorities "between individual patients and between patients from different cultures and health care systems", but there has been

little cross-country evidence to support these claims.³ The individual characteristics shown to be most closely associated with patient priorities in the literature are respondent age, followed by education, health status, sex, and utilization rates (age came up in 36% of studies).¹⁶ In a seminal study in 1991, the relative importance given to non-clinical quality dimensions was found to be similar for respondents utilizing both ambulatory and hospital inpatient services.¹⁷ Later studies have reported weak associations between rankings and individual level characteristics.^{16,18,19}

Few studies have had sufficient empirical reach to test these hypotheses across countries.⁴ The largest empirical study to do so recently covered only 12 European countries.²⁰ WHO's MCS Study has produced the first large dataset on the relative importance of several non-clinical dimensions of quality of care across a wide range of countries. Based on the literature, we expected to find that responsiveness respect-for-persons domains were very important to populations, and to see larger differences in domain importance across countries than within countries.^{12,21-24}

METHODS

Framing importance questions

Techniques for assessing importance, or priorities, range from complex preference trade-offs, popular in health economics²⁵, to simpler rating or ranking exercises, more common in quality of care literature, and also referred to as 'attitude' studies.³ Simple ranking methods were used in the Rheumatism QQuality Of care Through patients' Eyes (QUOTE) study⁷ and in the 12-country QUOTE study²¹, which used a rating on 0 to 10 scales.

Underlying questions on priorities is the theory of how people formulate their preferences. The theory of scarcity contends that an individual's "priorities reflect the socioeconomic environment" by placing "greater subjective value on those things in relatively short supply"²⁶ (page 220). Another theory that features prominently in the literature is described by van Campen, Sixma, et al.⁷, who relates importance judgments to attitudes and values, whose formation is connected to the process of socialization arising from cultural norms and sharing common institutions.²⁷

The MCS questionnaire, responsiveness module, and the importance questions

After field testing (n=811 in 8 countries and with 191 cognitive interviews), the importance question was developed for the responsiveness module in the MCS Study. It asked survey respondents to identify the most important domain, and the least important domain – in both cases, from a close-ended list of 8 domains. The questions took on average 5 minutes to administer (see Figure 5.1) and formed one of three parts to the responsiveness module. Other

parts covered user 'experiences', termed 'performance questions' (49 items) and one on within-domain standards (termed the 'expectations questions') (14 items) (see <http://www.who.int/responsiveness/surveys/en/>). The responsiveness module followed the same translation protocol administered for the questionnaire as a whole, involving forward and back translation of key terms by a third person and review of translations by a nationally appointed expert panel. Study questionnaires were translated into at least 1 and up to 3 national languages per country, and tested on at least 20 local respondents. Administration time for the responsiveness module was on average 20 minutes. The psychometric properties of the whole responsiveness module are published elsewhere.²⁸

MCS Study implementation

Surveys were subcontracted to principal investigators in countries, who aimed for national coverage (except in India, China and Nigeria where costs limited coverage to a few states). Sampling schemes used general population sampling frames and stratified multi-stage random sampling or cluster sampling with random walk. Within households, eligible respondents (18 years or older) were selected using the 'most recent birthday' method or Kish tables. Depending on the health modules included, the questionnaire took between 30 and 100 minutes to administer. The MCS Study protocol and processes were cleared with the WHO Sub-Committee for Research Involving Human Subjects.¹³

Survey respondents and non-respondents

Data extracted for analysis for this paper included 105,806 records from interviewer-administered interviews in 41 countries. Survey response rates were on average 79% for effective contacts and 46% for attempted contacts (see Table 5.1). Response rates were comparable to the Coulter and Jenkinson study²², and to a few studies in Sitzia and Wood¹⁴, they were lower than the median rate of 76% reported in Sitzia and Wood.¹⁴ Ex-post comparisons of the survey sample's age and sex profiles with UN population statistics showed that in both sexes, younger respondents (<35 years) were under-represented; older respondents (60-65 years) were over-represented.¹³ UN education statistics (averaging 8 years for the 41 countries) indicated biases towards more educated respondents.²⁹

Read the cards below. These provide descriptions of some different ways the health care services in your country show respect for people and make them the centre of care. Thinking about what is on these cards and about the whole health system, which is the most important and the least important to you?

DIGNITY

- being shown respect
- having physical examinations conducted in privacy

CONFIDENTIALITY OF INFORMATION

- having your medical history kept confidential
- having talks with health providers done so that other people who you don't want to have hear you can't overhear you

CHOICE

- being able to choose your doctor or nurse or other person usually providing your health care
- being able to go to another place for health care if you want to

PROMPT ATTENTION

- having a reasonable distance and travel time from your home to the health care provider
- getting fast care in emergencies
- short waiting times for appointments and consultations, and getting tests done quickly
- short waiting lists for non-emergency surgery

AUTONOMY

- being involved in deciding on your care or treatment if you want to
- having the provider ask your permission before starting treatments or tests

SURROUNDINGS OR ENVIRONMENT

- having enough space, seating and fresh air in the waiting room
- having a clean facility (including clean toilets)
- having healthy and edible food

SOCIAL SUPPORT

- being allowed the provision of food and other gifts by relatives while in hospital
- being allowed freedom of religious practices

COMMUNICATION

- having the provider listen to you carefully
- having the provider explain things so you can understand
- having time to ask questions

MOST IMPORTANT _____

LEAST IMPORTANT _____

Figure 5.1 The question on the importance of responsiveness domains (WHO MCS Study)

Responsiveness module and importance: inappropriate respondent and item 'non-response'

For the responsiveness module as a whole, 6% of respondents in the long surveys (12 countries) had inappropriate non-response for 5% or more items. For the remaining surveys, on average 15% of respondents had non-responses in more than 5% of items, and 7% had non-responses in more than 6% of items. Only 2% of respondents were missing both importance questions. Average item non-response across countries was 2% for the most important domain, and 10% for least important domain. Average missing rates for the most important domain were only higher than 5% in 4 (of 41) countries: Czech Republic (13%), Finland (6%), Mexico (12%), and Bulgaria (7%). For the least important domain missing rates were higher than 15% in 13 (of 41) countries (see Table 5.1). Based on the higher missing rates, results for the least important domain were anticipated to be less valid and reliable than results for the most important domain.

Analytical variables

Following the literature review on which individual level variables were important covariates with priority or preferences rankings, we extracted sex, age, educational status, visit to health services in the previous 12-months (hospital inpatient stays, ambulatory care only, or neither), and self-assessed health in the previous 30 days (very bad, bad, moderate, good, very good). Respondents rating their general health over the previous 30 days as good or very good were classified as 'healthy', and otherwise, as 'less healthy'. Among covariates, the education variable had the highest item missing (2.7%), but this rate was still acceptable.

Secondary data were also used to assess variations in responses at the country level. Given the emphasis on resources and cultural or institutional factors influencing attitudes and values in the literature, we selected variables for health expenditure per capita (adjusted for price parity and denominated in US Dollars), a population level human development indicator, and a geographic grouping of countries that corresponded roughly to cultural factors. Expenditure described the wealth of human and technological resources in the health service setting.³¹ The United Nations Development Programme's Human Development Index, a measure of overall population development, categorized countries as having either high or low (to medium) human development.³¹ Welzel, Inglehart, and Klingeman's²⁷ "cultural zones" variables split countries according to the following groupings: Western Europe (plus Canada), Latin America, Eastern Mediterranean (Islamic) countries, ex-Soviet, Asia, and Africa (Nigeria) (see Table 5.1). The few countries in Asia and Africa meant this geographic variable was not useful in these regions, but there was a good spread of countries across the other regions (see Table 5.1).

Analyses

This paper used frequency and multivariate analyses to describe the data and explore associations between individual level and country level characteristics and the selection of a responsiveness domain as most or least important. All analyses were performed using Stata Special Edition v7.

Frequency analyses

Standard frequency analyses were reported for each importance question by country and for socio-demographically defined population groups within countries (by sex, age, education, health status (self-reported health), and utilization) (Table 5.2). Results for population groups were also averaged across countries (Table available on request).

Multinomial logit regression

Multinomial logit regressions are generally used to estimate relationships where the response variable Y is measured in terms of $K+1$ categories. Our maximum likelihood multinomial logit regressions estimated, for each of the two questions on the most and least important domain, the probability of choosing domain x over a base domain, to which we assigned the most frequently selected domain – prompt attention in the case of the question on the 'most important domain', and access to social support networks (social support) in the case of the question on the 'least important domain'. Variables included in the logit models reflected features thought to be important in determining preferences, as discussed previously. Reference categories for categorical individual level variables were: males, "very bad" health, and having had a hospital inpatient stay in the previous 12 months. Age and education were continuous variables. For country level variables, expenditure per capita was continuous, human development was categorical – high human development was the reference category. Western Europe was the reference category for geographical groupings.

The regression results were presented for choices with respect to the most important domain in Table 5.3. The results focussed on listing the relative risk ratios, where p -values were less than 0.03 and the 95% confidence intervals excluded the relative risk of 1. Results for the least important regression model are reported in the text where they add to the analysis. The interpretation of the relative risk ratio is, holding all the other variables constant, the effect of a particular variable on the relative risk of choosing a particular domain over the base category domain (prompt attention or social support). The discussion of the regression results was aided by the use of thresholds of more than 1.5 or less than 0.5 to identify large influences on the probability of selecting a particular domain as most or least important. While these thresholds were an arbitrary rule-of-thumb borrowed from clinical epidemiology, they helped to distinguish larger from smaller deviations. The purpose was not to rule out the significance of certain explanatory

factors but to point to those factors with more influence. The discussion, in particular references to other results in the literature, was further aided by further analyses of choices with respect to other outcome categories in addition to the base reference category.

RESULTS

Characteristics of respondents

Across 41 countries, 105,806 respondents (52% female) completed the responsiveness module. Their average age was 42 years (SD=5 years). Their mean years of education was 11 years (SD=3 years). About half the respondents said their health was good or very good. About 56% (SD=13%) of respondents reported at least 1 visit to ambulatory services and 10% (SD=3%) reported at least one visit to a hospital in the previous 12 months (see Table 5.1 for country specific information).

Importance frequencies

For the MCS sample as a whole, the domain selected by Study respondents as most important was prompt attention (41%). Next came dignity (22%), communication (14%), choice (8%), confidentiality (6%), autonomy (4%), quality of basic amenities (3%), and access to social support (in hospital) (2%). Social support was most commonly selected as the least important domain (41%), followed by quality of basic amenities (13%), autonomy (12%), choice (12%), confidentiality (9%), and communication (5%). Prompt attention (4%) and dignity (3%) were least frequently selected as the least important domain, a finding which provided internal validation of the ranking observed in responses to the question on the most important domain. Averages across countries shown in Table 5.2 had the same ranking. In 37 of 41 countries, the highest percentage of respondents selected prompt attention as most important. The ranking of domains observed for subpopulations, where each country was weighted equally, showed a similar ordering.

Table 5.1 The WHO Multi-Country Survey Study Interviewer-administered surveys: sample descriptions for 41 countries

| Survey | Item missing rates: all items and 2 importance questions (%) | | | Final respondents n | Fe- males % | Age yrs | Edu- cation yrs | Response rate for attempted contacts % | Self- reported health (very good) % | At least 1 ambula- tory visit % | At least 1 hospital inpatient visit % |
|--------------------|--|------|-------|------------------------|-------------------|------------|-----------------------|--|--|--|---|
| | All | Most | Least | | | | | | | | |
| Belgium | 3 | 3 | 16 | 1,100 | 52 | 44 | 13 | 36(48) | 23 | 58 | 12 |
| Canada | 1 | 2 | 10 | 393 | 50 | 48 | 14 | 11(11) | 35 | 44 | 10 |
| Finland | 4 | 6 | 15 | 1,021 | 56 | 47 | 14 | 21(52) | 13 | 72 | 15 |
| France | 2 | 1 | 8 | 1,003 | 52 | 43 | 18 | 42(77) | 26 | 66 | 13 |
| Germany | 4 | 4 | 21 | 1,123 | 52 | 47 | 13 | 67(80) | 21 | 63 | 9 |
| Iceland | 6 | 2 | 13 | 489 | 54 | 39 | 20 | .(53) | 34 | 63 | 9 |
| Ireland | 3 | 1 | 17 | 711 | 50 | 42 | 12 | 17(39) | 45 | 48 | 13 |
| Italy | 3 | 2 | 19 | 1,002 | 52 | 45 | 12 | 36(61) | 13 | 46 | 7 |
| Luxembourg | 2 | 2 | 7 | 719 | 56 | 45 | 14 | .(55) | 25 | 73 | 13 |
| Malta | 1 | 2 | 26 | 500 | 51 | 47 | 12 | 48(59) | 13 | 65 | 10 |
| Netherlands | 4 | 2 | 23 | 1,085 | 53 | 38 | 7 | 55(59) | 23 | 39 | 8 |
| Portugal | 4 | 4 | 33 | 1,001 | 45 | 33 | 11 | 37(61) | 34 | 63 | 10 |
| Spain | 2 | 1 | 22 | 1,000 | 55 | 42 | 12 | 39(84) | 17 | 68 | 12 |
| Sweden | 3 | 1 | 6 | 1,000 | 51 | 43 | 11 | 19(75) | 18 | 62 | 9 |
| Argentina | 2 | 2 | 15 | 781 | 53 | 43 | 10 | 46(53) | 24 | 62 | 11 |
| Colombia | 3 | 2 | 2 | 6,019 | 65 | 40 | 7 | 72(84) | 12 | 64 | 9 |
| Costa Rica | 1 | 1 | 9 | 756 | 49 | 38 | 8 | 37(67) | 21 | 66 | 10 |
| Mexico | 6 | 12 | 16 | 4,812 | 40 | 42 | 9 | .(96) | 15 | 39 | 7 |
| Venezuela | 1 | 0 | 0 | 754 | 49 | 34 | 11 | .(66) | 31 | 36 | 7 |
| Bahrain | 1 | 0 | 0 | 809 | 45 | 35 | 11 | 35(44) | 56 | 50 | 10 |
| Egypt | 4 | 0 | 0 | 4,486 | 56 | 39 | 7 | .(99) | 38 | 59 | 6 |
| Iran | 0 | 1 | 2 | 9,568 | 52 | 38 | 7 | not available | 13 | 75 | 9 |
| Jordan | 3 | 0 | 0 | 803 | 51 | 34 | 10 | 74(83) | 49 | 50 | 12 |
| Morocco | 1 | 0 | 16 | 754 | 50 | 36 | 8 | .(69) | 13 | 58 | 7 |
| Oman | 3 | 0 | 0 | 835 | 61 | 36 | 8 | 67(79) | 45 | 21 | 3 |
| Syria | 0 | 1 | 1 | 8,624 | 54 | 48 | 10 | not available | 28 | 58 | 11 |
| Turkey | 4 | 4 | 5 | 5,197 | 43 | 33 | 10 | .(90) | 27 | 27 | 3 |
| U. Arab Emir. | 3 | 0 | 0 | 818 | 45 | 34 | 13 | 71(76) | 40 | 59 | 11 |
| Bulgaria | 0 | 7 | 28 | 1,000 | 51 | 45 | 14 | 69(88) | 16 | 59 | 8 |
| Croatia | 3 | 4 | 8 | 1,500 | 63 | 48 | 11 | .(68) | 19 | 57 | 11 |
| Czech Rep. | 4 | 13 | 31 | 1,072 | 40 | 44 | 14 | 55(60) | 18 | 68 | 14 |
| Estonia | 6 | 4 | 19 | 1,000 | 57 | 48 | 10 | 50(71) | 7 | 73 | 15 |
| Georgia | 1 | 0 | 0 | 9,847 | 58 | 46 | 12 | 84(93) | 13 | 23 | 4 |
| Latvia | 2 | 3 | 12 | 752 | 49 | 48 | 12 | 54(72) | 3 | 61 | 16 |
| Romania | 5 | 4 | 15 | 1,051 | 56 | 45 | 9 | 39(52) | 9 | 63 | 9 |
| Russia | 3 | 2 | 13 | 1,601 | 50 | 45 | 14 | 25(10) | 12 | 52 | 16 |
| Slovakia | 3 | 2 | 3 | 1,183 | 54 | 43 | 15 | 39(84) | 6 | 60 | 13 |
| China ^a | 1 | 0 | 1 | 9,442 | 47 | 40 | 9 | .(99) | 29 | 48 | 8 |

Table 5.1 The WHO Multi-Country Survey Study Interviewer-administered surveys: sample descriptions for 41 countries (continued)

| Survey | Item missing rates: all items and 2 importance questions (%) | | Final respondents | Fe-males | Age | Edu-cation | Response rate for attempted contacts | Self-reported health (very good) | At least 1 ambulatory visit | At least 1 hospital inpatient visit | |
|----------------------|--|---|-------------------|----------|-----|------------|--------------------------------------|----------------------------------|-----------------------------|-------------------------------------|---|
| India ^a | 1 | 1 | 1 | 5,196 | 53 | 40 | 4 | (.98) | 15 | 63 | 8 |
| Indonesia | 1 | 1 | 2 | 9,952 | 55 | 40 | 7 | (.99) | 13 | 41 | 3 |
| Nigeria ^a | 1 | 1 | 4 | 5,047 | 55 | 44 | 14 | 45(.98) | 19 | 63 | 8 |

^aSub-national samples: China (Shandon, Henan and Gansu); India (Andhra Pradesh); Nigeria (Oyo State)

Table 5.2 Percentage of respondents selecting domains as most important (n=41)

| Country | Dignity | Autonomy | Communi-cation | Confiden-tiality | Choice | Prompt attention | Basic Amenities | Social support |
|----------------|-----------|----------|----------------|------------------|-----------|------------------|-----------------|----------------|
| Belgium | 18 | 7 | 10 | 14 | 13 | 37 | 1 | 1 |
| Canada | 7 | 3 | 13 | 12 | 8 | 51 | 1 | 4 |
| Finland | 16 | 4 | 5 | 8 | 6 | 61 | 0 | 0 |
| France | 12 | 5 | 13 | 13 | 13 | 42 | 1 | 1 |
| Germany | 11 | 8 | 5 | 19 | 16 | 38 | 0 | 1 |
| Iceland | 24 | 7 | 11 | 21 | 8 | 28 | 1 | 0 |
| Ireland | 19 | 3 | 7 | 10 | 16 | 44 | 1 | 1 |
| Italy | 9 | 3 | 6 | 5 | 10 | 64 | 2 | 1 |
| Luxembourg | 15 | 5 | 15 | 17 | 12 | 31 | 2 | 2 |
| Malta | 12 | 8 | 14 | 14 | 11 | 37 | 2 | 1 |
| Netherlands | 18 | 7 | 12 | 14 | 7 | 40 | 1 | 1 |
| Portugal | 10 | 3 | 6 | 6 | 15 | 57 | 2 | 1 |
| Spain | 10 | 4 | 13 | 5 | 7 | 59 | 1 | 1 |
| Sweden | 20 | 8 | 12 | 4 | 3 | 53 | 0 | 0 |
| Average | 14 | 5 | 10 | 12 | 10 | 46 | 1 | 1 |
| Argentina | 22 | 2 | 8 | 4 | 12 | 48 | 2 | 1 |
| Colombia | 20 | 5 | 15 | 4 | 9 | 45 | 2 | 1 |
| Costa Rica | 17 | 3 | 19 | 6 | 10 | 39 | 3 | 3 |
| Mexico | 21 | 5 | 11 | 7 | 5 | 46 | 3 | 3 |
| Venezuela | 21 | 2 | 6 | 5 | 5 | 58 | 2 | 3 |
| Average | 20 | 3 | 12 | 5 | 8 | 47 | 2 | 2 |
| Bahrain | 33 | 2 | 19 | 7 | 6 | 29 | 2 | 2 |
| Egypt | 49 | 3 | 10 | 9 | 6 | 19 | 4 | 1 |

**Table 5.2 Percentage of respondents selecting domains as most important (n=41)
(continued)**

| Country | Dignity | Autonomy | Communication | Confidentiality | Choice | Prompt attention | Basic Amenities | Social support |
|----------------------|-----------|----------|---------------|-----------------|-----------|------------------|-----------------|----------------|
| Iran | 21 | 3 | 16 | 13 | 5 | 31 | 8 | 3 |
| Jordan | 36 | 3 | 9 | 5 | 4 | 40 | 2 | 0 |
| Morocco | 41 | 1 | 12 | 2 | 3 | 38 | 3 | 1 |
| Oman | 25 | 3 | 20 | 9 | 8 | 31 | 3 | 1 |
| Syria | 53 | 3 | 8 | 6 | 4 | 22 | 2 | 2 |
| Turkey | 17 | 4 | 13 | 3 | 5 | 46 | 8 | 4 |
| U.Ar.Emir. | 20 | 3 | 16 | 9 | 8 | 40 | 3 | 2 |
| Average | 33 | 3 | 14 | 7 | 5 | 33 | 4 | 2 |
| Bulgaria | 10 | 1 | 3 | 5 | 17 | 63 | 1 | 1 |
| Croatia | 17 | 2 | 14 | 4 | 9 | 51 | 1 | 1 |
| Czech Rep | 11 | 6 | 4 | 6 | 10 | 61 | 2 | 0 |
| Estonia | 13 | 1 | 8 | 7 | 25 | 42 | 1 | 3 |
| Georgia | 25 | 3 | 21 | 4 | 13 | 26 | 5 | 3 |
| Latvia | 19 | 1 | 11 | 5 | 20 | 38 | 1 | 4 |
| Romania | 16 | 1 | 10 | 4 | 5 | 58 | 3 | 2 |
| Russia | 12 | 3 | 9 | 6 | 12 | 57 | 1 | 1 |
| Slovakia | 13 | 6 | 19 | 14 | 10 | 36 | 1 | 2 |
| Average | 15 | 3 | 11 | 6 | 13 | 48 | 2 | 2 |
| China | 20 | 7 | 14 | 6 | 10 | 37 | 3 | 3 |
| India | 14 | 1 | 23 | 2 | 3 | 53 | 2 | 3 |
| Indonesia | 16 | 3 | 12 | 2 | 5 | 61 | 2 | 1 |
| Average | 17 | 4 | 16 | 3 | 6 | 50 | 2 | 2 |
| Nigeria | 7 | 2 | 21 | 4 | 2 | 58 | 5 | 2 |
| Average (all) | 19 | 4 | 12 | 8 | 9 | 44 | 2 | 2 |

Multinomial logit regressions

The Pseudo R-square for the model of the most important domain was 0.0327 and the Chi-square value was 11,155. Given the size of the dataset, the Chi-square value was significant as expected ($p < 0.0000$). Relative risk ratios were calculated for all covariates with respect to each of the domains and the reference category.

Table 5.3 lists shows how most risk ratios for the analysis of the most important domain ranged between 0.5 and 1.5 (including the extreme categories for the 95 % confidence interval). Most of the more extreme relative risk ratios¹⁹ were reported for geographical groupings. Excluding the 'Africa' grouping, which only contained 1 country survey, the number of relative risk ratios was 14. For example as seen in the last row and third column of Table 5.3, there was a relatively higher risk of choosing confidentiality as the most important domain in Eastern Mediterranean (relative risk ratio of 3.68). The relative risk of choosing basic amenities and social support as the most important domain was higher in all regions relative to Western Europe. In countries with low (to medium) human development, the elevated relative risks of choosing basic amenities over prompt attention as the most important domain was 2.64 (95% confidence interval of 2.24 - 3.11). Individual-level variables with a marked impact on choices of priorities were health status (in all cases except for basic amenities) and utilization of health services (5 out of 7 domains - not autonomy and choice), but these relative risk ratios did not exceed the threshold values. The least important domain regression model had a Pseudo R-square of 0.0246 and a Chi-square value of 8,641 (n=100,305). Results largely mirrored those in the most important domain regression.

Table 5.3 Relative risks of selecting a domain over “prompt attention” as most important: significant relative risks [with 95% confidence intervals; p<0.03; n= 103,081] from the multinomial logistic regression modela (reference category comparison)

| | 1 (dignity) | 2 (confidentiality) | 3 (choice) | 5 (autonomy) | 6 (basic amenities) | 7 (social support) | 8 (communication) |
|---|------------------------------|--|------------------------------|------------------------------|---------------------------|--|--|
| Sex ^b | 1.11 [1.07-1.15] | 1.11 [1.05-1.17] | none | none | none | none | 1.10 [1.05-1.14] |
| Age | 0.99 [0.99-0.99] | none | none | 1.03 [1.02-1.03] | none | none | none |
| Education | None | 0.99 [0.99-0.99] | 1.02 [1.01-1.02] | none | none | none | none |
| Health ^b | "very good" 1.24 [1.08-1.43] | "moderate" 1.51 [1.15-1.98] "good" 1.45 [1.10-1.91] "very good" 1.80 [1.37-2.38] | "very good" 1.25 [1.03-1.52] | "very good" 1.68 [1.19-2.36] | none | "bad" 0.66 [0.48-0.93] "moderate" 0.68 [0.50-0.91] "good" 0.55 [0.40-0.74] "very good" 0.66 [0.48-0.91] | "good" 0.84 [0.73-0.98] |
| Utilization (last 12 months) ^b | "no use" 1.12 [1.05-1.20] | "ambulatory" 1.15 [1.03-1.27] "no use" 1.23 [1.10-1.37] | none | none | "no use" 1.33 [1.14-1.54] | "ambulatory" 0.80 [0.67-0.96] "no use" 1.36 [1.14-1.63] | "ambulatory" 1.25 [1.16-1.36] "no use" 1.15 [1.06-1.24] |
| Human development ^b | 1.06 [1.09-1.24] | none | 1.25 [1.15-1.26] | 1.30 [1.15-1.47] | 2.64 [2.24-3.11] | none | 1.37 [1.28-1.48] |

Table 5.3 Relative risks of selecting a domain over “prompt attention” as most important: significant relative risks [with 95% confidence intervals; $p < 0.03$; $n = 103,081$] from the multinomial logistic regression modela (reference category comparison) (continued)

| | 1 (dignity) | 2 (confidentiality) | 3 (choice) | 5 (autonomy) | 6 (basic amenities) | 7 (social support) | 8 (communication) |
|-------------------------------------|--------------------------------|--|---------------------------------|--------------------------------|--|--|--|
| Health expenditure per capita | 0.93 [0.92-0.94] | 1.07 [1.06-1.08] | 1.03 [1.02-1.04] | 1.05 [1.03-1.06] | 1.09 [1.06-1.12] | 1.03 [1.01-1.06] | |
| Geographical groupings ^b | Latin America 0.47 [0.41-0.54] | Eastern Mediterranean 3.68 [2.97-4.57] | Eastern Europe 2.20 [1.81-2.66] | Latin America 1.55 [1.20-2.00] | Latin America 5.05 [3.32-7.69] | Latin America 2.56 [1.65-3.95] | Latin America 1.38 [1.17-1.63] |
| | Ex-Soviet 0.45 [0.39-0.52] | Ex-Soviet 2.03 [1.65-2.50] | Africa 0.24 [0.18-0.32] | Ex-Soviet 1.79 [1.34-2.4] | Eastern Mediterranean 15.71 [9.86-25.04] | Eastern Mediterranean 5.27 [3.26-8.51] | Eastern Mediterranean 1.87 [1.55-2.25] |
| | Asia 0.24 [0.21-0.28] | | | Ex-Soviet 1.51 [1.11-2.06] | Ex-Soviet 9.47 [5.98-15.00] | Ex-Soviet 4.11 [2.58-6.55] | Ex-Soviet 2.00 [1.67-2.39] |
| | Africa 0.07 [0.06-0.09] | | | Africa 0.44 [0.30-0.65] | Asia 4.87 [2.97-7.98] | Asia 2.66 [1.60-4.42] | Asia 1.38 [1.14-1.68] |
| | | | | | Africa 7.98 [4.77-13.70] | Africa 2.24 [1.287-3.93] | Africa 1.59 [1.28-1.97] |

^a Model statistics: LR $\chi^2(112) = 11154.53$ (Prob > $\chi^2 = 0.0000$), Log likelihood = -164834.93, Pseudo R² = 0.0327;

^b Dummy variables; base categories: sex - male; health - “very bad”; utilization - hospitalization; human development - high; geographical groupings - Western Europe

DISCUSSION

We compared the relative importance of responsiveness domains across 41 countries and across different subpopulations within countries. Prompt attention, dignity and communication were most frequently identified as most important, and least frequently identified as least important. Across subpopulations within countries, convergence was stronger than convergence across countries, indicating that health system investments, culture and the human development context were stronger influences on populations' priorities for their health systems than individual level factors like age, sex, education, health status, and utilization of health services. The similarity of the domain rankings, whether ranked by the most or least important question strongly supported the consistency of the results. A further validity check was provided by the WHO pilot surveys, which used a full ranking question. The system of two ranking questions from the MCS Study obtained the same domain ranking as the full ranking question from the pilot surveys.

Comparing results with other studies

Results from our analysis of the WHO responsiveness data largely agreed with results from other studies. Technical quality, which is closely related to the responsiveness domain of prompt attention (and improving health outcomes), was considered more important than non-technical quality in Fung et al.¹⁸ Dignity, the responsiveness domain most closely related to humaneness of treatment, was also as strong contender for first place, where it has been placed in a number of other studies. The probable explanation for the greater emphasis on prompt attention in the WHO surveys was that unlike other studies, the WHO questionnaire operationalized the concept of prompt attention in terms of geographical access and access in case of emergencies (see Figure 5.1). Other surveys incorporating this dimension usually focused on waiting times.^{3,4} Also, the WHO importance question asked about the whole health system, including inpatient or hospital care; most other studies focused on ambulatory services. Accessibility to inpatient services, especially for serious health crises is very important to populations, as was found in a five-country survey of populations' concerns with health systems (USA, Canada, UK, New Zealand, Australia), which covered issues related to costs, types of services, shortages, waiting times, and government funding. In this study, waiting times for surgery and treatment for more serious conditions was highlighted as one of the biggest challenges facing health care systems in 3 of out 5 countries.³²

Despite observing overall convergence of importance rankings, we found a few differences between population groups within countries and several more instances of differences across countries. Across population groups described by socio-demographic characteristics, our strongest results emerging from the logit regressions were for divergences according to an individual's health status. For country-level variables, we found that choices (or preferences) were mostly influenced by the country's geographic zone and level of human development, although in some instances - notably for basic amenities - the level of health expenditure in the country was an important determinant of priorities.

Regression results for individual level characteristics revealed several more relationships, some of which compared well with results from other studies. For example, Blendon, Schoen, et al.³² have argued that reports from 'the sick', defined as health system users with experiences of inpatient stays within the previous 2 years (other than for normal, uncomplicated deliveries) and who rated their health as poor, have greater validity. Our results also revealed differences in the selection of domains according to health status. Differences in the importance of domains within countries were also related to the respondent's utilization of the system (13 out of 41 countries and in 5 out of 7 domains), but, in general, there was support for convergence in service priorities across user groups as found in another study for patients from general practice, dental, and hospital settings.¹⁷ Some results differed from those in other studies: we found that

older respondents considered patient autonomy as slightly more important than did younger respondents.^{33,34} The Coulter and Jenkinson²² study of 8 European countries found that although the majority of people preferred the shared decision-making model (51%), 31% of those aged over 55 were likely to say that the doctor should decide, compared with 24% under the age of 35. We had observed similar trends in the raw frequency analyses for similar age cohorts, but the associations reversed in the multivariate regression analysis when age was entered as a continuous variable. One reason for this may be that the relationship applies to cohorts rather than age and may be non-linear.²⁶

Certain divergences in priorities we observed between countries may be explained by the theory of scarcity. In our study, countries with low human development and low levels of health expenditure placed relatively higher importance on basic amenities. Another example is the high ranking of “choice” by ex-Soviet countries, which probably relates to the previous limitations on choice experienced by these populations prior to the demise of communism.

Sampling

We are of the opinion that the convergence we found in the importance of domains was not undermined by either the low response rates or by the over-representation of educated populations observed for some countries. Though we expect that specific priority ranking within countries may differ slightly with improved sampling accuracy. Clearly, the data cannot have done justice to the diversity of the populations in China and India, but this was all the data available. There are several studies within countries in Asia and Africa measuring quality performance in these dimensions, but much fewer, to our knowledge, that focus on importance.^{35,36}

Study design weaknesses

Several weaknesses in the study design merit discussion. First, instead of complex preference trade-off techniques³⁷, the questionnaire used a simple and straightforward ranking exercise (as did the European Task Force on Patient Evaluations of General Practice (EUROPEP)).³⁸ Asking only about the most and least important domains means that less is known about ranks two to seven. Second, given the close-ended nature of the questions, the results do not rule out the possibility of there being a domain more or less important than the ones listed. Literature related to the concept of responsiveness have suggested that issues like empathy, warmth and coordination of care may be missing from the existing responsiveness domain structure.^{39,40} Coordination of care is a product of good communication between care providers and prompt attention, and is another characteristics that would probably have been rated highly across countries, judging from other studies.³²

Third, the low association between importance rankings and individual characteristics may be partly explained by the omission of individual characteristics like income and ethnicity. Studies in psychiatric care have found ethnicity to be an important determinant of preferences.⁴¹ Our study did include the most commonly included socio-demographic variables studies in the literature and widely available country-level variables, although an in-depth study to generate specific socio-demographic variables of interest in a given context and variables related to the design of health institutions and empowerment would be desirable.⁴²

Implications for policy-making

Our results have two broad implications for policy-making. First, responsiveness is a concept encompassing dimensions that resonate with what individuals value and want in their health care. Combining importance and performance information into a single summary measure, using a common set of weights, may be envisaged as a means of monitoring quality of care from the user's perspective within countries. But despite general convergence, there is still sufficient divergence to remain prudent about comparisons using summary measures, especially international comparisons. We propose that international comparisons should be confined to specific domains, which individuals from a variety of cultures and contexts are apparently able to distinguish between. This conservative approach is currently envisaged for the Organisation for Economic Cooperation and Development (OECD) Health Care Quality Indicators Project.⁴³

Second, policy makers can use these results from these surveys to prioritize efforts when resources are limited. As observed elsewhere, importance ratings are more likely than performance ratings to be more stable across countries over time.²¹ The data provide a clear message to prioritize reforms that improve prompt attention, but not at the expense of patient dignity and communication, which may damage the acceptability of health services to users, and result in barriers to access. When comparing health systems performance within and across countries, policymakers are usually interested in performance ratings, but efforts to remedy performance weaknesses, which are frequently aimed at decreasing costs and improving supply-side efficiency, may be misguided without a better understanding of population priorities. Designing appropriate mechanisms that permit population priorities to shape health system reforms represents another policy-relevant field of enquiry for improving the user-responsiveness of health systems.^{44,45}

REFERENCES

1. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
2. Campbell SM, Roland MO, Buetow SA. Defining quality of care. *Soc Sci Med.* 2000;51(11):1611–25.
3. Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. *Soc Sci Med.* 1998;47(10):1573–88.
4. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med.* 1997;45(12):1829–43.
5. Wensing M, Elwyn G. Methods for incorporating patients' views in health care. *Brit Med J.* 2003;326(7394):877–9.
6. Cleary PD. The increasing importance of patient surveys. *Brit Med J.* 1999; 319(720):721.
7. van Campen C, Sixma JJ, Kerssens HJ, Peters L, Rasker JJ. Assessing patients' priorities and perceptions of the quality of health care: the development of the QUOTE rheumatic-patients instrument. *Brit J Rheumatol.* 1998;37(4):362–8.
8. World Health Organization. Health systems: improving performance. The world health report 2000. Geneva: World Health Organization; 2000.
9. Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bull World Health Organ.* 2000;78(6):717–31.
10. Gostin L, Hodge J, Valentine N, Nygren-Krug H. The domains of health responsiveness: a human rights analysis. Health and human rights working paper no. 2. Geneva: World Health Organization; 2003. [online: http://apps.who.int/iris/bitstream/10665/73926/1/HHRWPS2_eng.pdf]
11. de Silva A. A framework for measuring responsiveness. Global programme on evidence for health policy discussion paper no. 32. Geneva: World Health Organization; 2000.
12. Mooney G, Wiseman V. World health report 2000: challenging a world view. *J Health Res Policy Syst.* 2000;5:198–9.
13. Üstün TB, Chatterji S, Villanueva M, Bendib L, Celik C, Sadana R, et al. WHO Multi-Country Survey Study on Health and Responsiveness 2000–2001. In Murray CJL, Evans DB, editors. Health systems performance assessment. Geneva:World Health Organization; 2003. p. 761–96.
14. Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care.* 1998;10(4):311–7.
15. Dillman DA, Tarnai J. Mode effects of cognitively designed recall questions: a comparison of answers to telephone and mail surveys. In Biemer P, Groves RM, Lyberg LE, Mathiowetz NA, Sudman S, editors. New York:Wiley-Interscience Publication; 1991. p. 73–93.
16. Jung HP, Baerveldt C, Olesen F, Grol R, Wensing M. Patient characteristics as predictors of primary health care preferences: a systematic literature analysis. *Health Expect.* 2003;6(2):160–81.
17. Williams S, Calnan M. Convergence and divergence: assessing criteria of consumer satisfaction across general practice, dental and hospital care settings. *Soc Sci Med.* 1991;33(6):707–16.

18. Fung CH, Elliott MN, Hays RD, Kahn KL, Kanouse DE, McGlynn EA et al. Patients' preferences for technical versus interpersonal quality when selecting a primary care physician. *Heath Serv Res.* 2004;40(4):957–77.
19. Sofaer S, Crofton C, Goldstein E, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? *Health Serv Res.* 2005;40(6p2):2018–37.
20. Kerssens JJ, Groenewegen PP, Sixma HJ, Boerma WGW, van der Eijk I. Comparison of patient evaluations of health care quality in relation to WHO measures of achievement in 12 European countries. *B World Health Organ.* 2004;82(2):106–14.
21. Groenewegen PP, Kerssens JJ, Sixma HJ, van der Eijk I, Boerma WGW. What is important in evaluating health care quality? An international comparison of user views. *BMC Health Serv Res.* 2005;5(16):1–9.
22. Coulter A, Jenkinson C. European patients' views on the responsiveness of health systems and healthcare providers. *Eur J Public Health.* 2005;15(4):355–60.
23. Hakkinen U, Ollila E, Taipale V, Koivusalo U, Rissanen P, Sintonen H. The world health report 2000: what does it tell us about health systems? Analyses by Finnish experts in Themes from Finland 7/2000. Helsinki:National Research and Development, Center for Welfare and Health (STAKES); 2000.
24. Gandek B, Ware JE, Jr., Aaronson NK, Alonso J, Apolone G, Bjorner J, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. *International Quality of Life Assessment.* *J Clin Epidemiol.* 1998;51(11):1149–58.
25. Ryan M. Discrete choice experiments in health care. *Brit Med J.* 2004;328(7436):360–1.
26. Inglehart R. Globalization and postmodern values. *Wash Q.* 2000(winter);23(1):215–28.
27. Welzel C, Inglehart R, Klingemann H. The theory of human development: a cross-cultural analysis. *Eur J Polit Res.* 2003;42: 341–79.
28. Valentine NB, Bonsel GJ, Murray CJL. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Qual Life Res.* 2007;16(7):1107–25.
29. United Nations Educational, Scientific and Cultural Organization [Internet]. Paris: UNESCO Data Centre, 'Education' statistics [cited May 2005]. Available from: <http://www.uis.unesco.org/>
30. World Health Organization (WHO). Reducing risks, promoting healthy life. The world health report 2002. Geneva; WHO; 2002.
31. United Nations Development Agency (UNDP). Human development report. UNDP: New York; 2001.
32. Blendon RJ, Schoen C, DesRoches CM, Osborn R, Zaper K. Common concerns amid diverse systems: health care experiences in five countries. *Health Affair.* 2003(May-Jun);22(3):106–21.
33. Thompson SC, Pitts JS, Schwankovsky L. Preferences for involvement in medical decision-making: situational and demographic influences. *Patient Educ Couns.* 1993 (Dec 31);22(3):133–40.
34. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). *Soc Sci Med.* 1997;44(5):681–92.

Chapter 5. Health System Responsiveness by N Valentine

35. Sekimoto M., Asai A, Ohnishi M, Nishigaki E, Fukui T, Shimbo T et al. Patients' preferences for involvement in treatment decision-making in Japan. *BMC Fam Pract.* 2004(Mar);5(1)e1-10. doi:10.1186/1471-2296-5-1.
36. Westaway MS, Rheeder P, Van Zyl DG, Seager JR. Interpersonal and organizational dimensions of patient satisfaction: the moderating effects of health status. *Int J Qual Health C.* 2003;15(4): 337–44.
37. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ.* 2002;21(2): 271–92.
38. Grol R, Wensing M, Mainz J, Ferreira P, Hearnshaw H, Hjortdahl P et al. Patient priorities with respect to general practice care: an international comparison. *European Task Force on Patient Evaluations of General Practice (EUROPEP). Fam Pract.* 1999;15(1): 4–11.
39. Stewart M. Reflections on the doctor–patient relationship from evidence and experience. *Brit J Gen Pract.* 2005;55(519):793–801.
40. Hudson S, Weisman C, Anderson R, Camacho F. The development and validation of the primary care satisfaction survey for women. *Women Health Iss.* 2004(Mar-Apr);14(2):35–50.
41. Fischer E, Shumway M, Owen R. Priorities of consumers, providers and family members in the treatment of schizophrenia. *Psych Serv.* 2002;53(6):724–9.
42. Atkinson S. Political cultures, health systems and health policy. *Soc Sci Med.* 2002;55(1):113–24.
43. Kelley E, Hurst J. Health care quality indicators project conceptual framework paper. Organisation for Economic Co-operation and Development (OECD) Working Papers 2006. Paris: OECD; 2006.
44. Wait S, Nolte E. Public involvement policies in health: exploring their conceptual basis. *Health Econ Policy Law.* 2006;1(2):149–62.
45. Pickard S, Sheaff R, Dowling B. Exit, voice, governance and user-responsiveness: the case of English primary care trusts. *Soc Sci Med.* 2006;63(2):373–83.

CHAPTER 6

What explains users' reports on health system responsiveness? Exploring health-care service and personal characteristics from surveys in 49 countries



N.B. Valentine, G.J. Bonsel
Manuscript submitted for publication

ABSTRACT

Introduction. Health care quality improvement is a common focus of national health policies. Yet few studies have described how user evaluations of quality are impacted by users' personal characteristics versus specific health-care service characteristics .

Methods. The WHO health system multi-domain responsiveness concept is used to study health-care quality from the user's perspective, drawing on data of 118,338 respondents in 49 countries from the published World Health Survey cross-sectional datasets (2002-2004). Multilevel fixed random intercept probit regression analyses estimate the association of responsiveness domains with 20 individual-level variables: 12 health service and 8 users' personal characteristics, with the country-level modelled by a country dummy.

Results. Regression coefficients for service characteristics range from 0.06 to 0.62 (absolute values, p -value <0.05), compared with 0.05 to 0.15 for users' personal characteristics. Discrimination, longer travel time, and management by the government is associated with worse responsiveness. Government management impacts mostly choice and basic amenities domains; whereas discrimination impacts mostly dignity.

Discussion. Health system responsiveness measures modifiable aspects of healthcare service quality that are likely to be impacted by specific health policy reforms.

Conclusion. Policy-makers should measure responsiveness to understand the impacts of their policies and to promote patient-centred quality of care in progress towards universal health coverage as part of the Sustainable Development Goals.

INTRODUCTION

Progressing toward universal health coverage (UHC) is a target under Goal 3 of the Sustainable Development Goals (SDGs). Monitoring UHC uses two indicators: the effective coverage of populations with a specific set of health interventions and coverage with financial health protection.¹⁻³ What this measurement approach is missing are measures of health-care quality from the users' perspective. Indicators for evaluating service qualities can be valued as ends in themselves as well as for facilitating more effective care. For example, having clear communication regarding medical diagnosis can be valued for peace of mind, but health workers with improved communication skills can also improve patient adherence to treatment. Monitoring these qualities of health systems is possible using existing, reliable measures.¹

Health system responsiveness is a concept described by the World Health Organization (WHO) as the quality of users' experiences with respect to a set of non-clinical attributes of health services, or domains. The responsiveness concept builds on fields of research dedicated to process and structural quality of health care⁴⁻⁸, and to health service access⁹. In the responsiveness concept, domains of prompt attention, choice, quality of basic amenities and social support are described as "client orientation" domains and assumed to be improved through increased material inputs. The other domains - dignity, communication, autonomy and confidentiality, are described as "respect for persons" domains, and associated with improved human rights mechanisms and codes of professional conduct.⁵ These research fields acknowledge that responsiveness attributes underpin good quality health services: intrinsically, because people value them^{10,4}, and instrumentally too, as facilitators of effective health service utilization and coverage.^{11,12} An example of the instrumental value is the evidence showing that poor responsiveness discourages people from seeking care or from completing hospital stays.^{13,14}

The WHO approach to measuring responsiveness differed from some previous approaches used to assess non-clinical quality of health-care services. For example, Donabedian's seminal work typically focused on the measurement of inputs^{7,8}, whereas responsiveness takes the experiences of users (patients, clients) as the point of departure, shifting measurement of quality of care towards an outcome. An attribute like respect for patient autonomy becomes judged by statements of users who have asked for medical advice rather than by characteristics of hospital organisation or the presence of guidelines for professional conduct. The advantage of this approach is that coverage of the population with quality services can be estimated on a common scale. But in order for responsiveness measurement to be useful in designing and reforming health systems, it is important to understand the link between common policy reforms (e.g., privatisation, purchasing integrated services, in-service training), their intermediate consequences

(e.g., type of health worker seen, access times, interpersonal relations) and the impact on measures of responsiveness.

This study is the first comprehensive research that addresses the last part of this causal chain. All responsiveness domain outcomes are studied as dependent variables using regression analyses applied to the WHO dataset from a cross-sectional, multi-country survey, WHO's World Health Survey (WHS) (2002-2004).¹⁵ The regression analyses explore how responsiveness is explained by a comprehensive set of intermediate factors, which include health-care service characteristics, as observed by users, and users' personal characteristics. Several studies have used the WHS data to describe the individual and country-level characteristics that drive 'differential reporting' behaviour in responsiveness and hence comparability of scores between users.¹⁶⁻²⁰ Yet the main question of what explains true performance in the responsiveness scores has received less attention in the literature. This study complements a small number of other studies that have researched macro, country-level determinants of responsiveness, such as country expenditure or national culture, the first part in the causal chain.¹⁷

This study is also the largest yet undertaken, covering 49 country datasets. Other studies have focussed on single country datasets (e.g., South Africa) or fewer characteristics explaining responsiveness, more often using only personal characteristics.^{21, 22} While using personal characteristics to explain patterns in responsiveness is important, in particular for understanding disparities or inequities, these analyses do not explain how health service characteristics potentially impact responsiveness. Addressing this latter question has greater scope for informing the use of responsiveness measures in tracking progress toward UHC as it will illuminate which factors are most likely to impact responsiveness when undertaking health system reforms.

METHODS

Data source

The WHS is a cross-sectional household survey programme that was launched by WHO in 2002.¹⁵ Its survey instruments were eventually adapted and reused for the extant longitudinal Study on global AGEing and adult health (SAGE), that covers the over-50 year old adult populations in a few countries (<http://www.who.int/healthinfo/sage/en/>). The WHS promoted the global collection of comparable health systems information, including - for the first time - responsiveness data (<http://www.who.int/healthinfo/survey/en/>). Between 2002 and 2004, WHS questionnaires were administered through household-based surveys in 71 countries by national and multinational contractors (e.g., GALLUP) under the overall coordination of the WHO. Ethical approval was obtained from an independent ethics review conducted by the Harvard School of Public Health's Institutional Review Board.²³ This paper analyses data from 49 out of the 57

WHS published survey datasets that used the long questionnaire, which contained an extensive set of questions characterising the health services. The 49 countries were selected on the basis of their use of the long questionnaire as well as their sampling frame completeness. The final dataset consisted of 46,477 inpatient and 76,426 outpatient records (Table 6.1). Several studies have published findings describing the adequacy of the WHS questionnaire's quality and its cross-country psychometric properties.^{15,18,24}

Responsiveness (outcome) measures or scores

The 15 dependent responsiveness variables for the regression analyses were extracted from the WHS dataset. The WHS questionnaire contained 7 responsiveness questions for outpatient (ambulatory) visits and 8 responsiveness questions for inpatient visits (Figure 6.1). The same domains (autonomy, communication, confidentiality, dignity, prompt attention, quality basic amenities and choice) are covered for outpatient and inpatient visits. But social support, was a domain covered only covered in inpatient visits. Ordinal responses to the 15 responsiveness questions were on a 5-point response scale ranging from "very good" to "very bad". For this study, individuals' responses on each domain question are dichotomized into two categories: "very good" and "good" (0), and "moderate", "bad" and "very bad" (1). The term responsiveness 'problem' (1=yes/0=no) refers to the dichotomized response variable.

Dichotomizing the dependent variable is a simple way to address individual-level differential reporting behaviour bias, which threatens commensurability of self-reported measures like self-assessed health (SAH) and responsiveness. Another approach is to use statistical models to develop continuous scores. Both approaches address the challenge of separating the measure of true responsiveness from the effects of differing expectations or respondent use of language and its impact on the use of the response scale. The dichotomization approach broadly conceptualizes reporting behaviour bias as shifts towards the "very good" or "very bad" ends of the ordinal scale, which have in fact commonly been found to exist.²⁰ Due to the high computational demands, using other more complicated statistical modelling for adjustment (for example, anchoring vignettes²⁵) would have restricted the analyses to roughly three predictor variables, which would have been too restrictive for the exploratory work aimed for in this study. Dichotomization presented a reasonable alternative, being used previously in studies exploring the determinants of SAH, which is also subject to reporting scale bias.²⁶

Table 6.1 Country samples (alphabetic): total, outpatients (12 months), inpatients (5 years) and share of total dataset for each

| Country name | Country Level of Income | Sample | Use of outpatient services | Percent of pooled data (outpatients) | Use of inpatient services | Percent of pooled data (inpatients) |
|------------------------|-------------------------|--------|----------------------------|--------------------------------------|---------------------------|-------------------------------------|
| Bangladesh | 1 | 5,065 | 4,064 | 5.3 | 1,007 | 2.1 |
| Bosnia and Herzegovina | 2 | 729 | 413 | 0.5 | 333 | 0.7 |
| Brazil | 3 | 3,966 | 2,343 | 3.1 | 1,623 | 3.4 |
| Burkina Faso | 1 | 1,999 | 1,301 | 1.7 | 815 | 1.7 |
| Chad | 1 | 1,084 | 602 | 0.8 | 624 | 1.3 |
| China | 2 | 2,006 | 1,480 | 1.9 | 601 | 1.3 |
| Comoros | 1 | 1,002 | 582 | 0.8 | 490 | 1 |
| Congo | 1 | 1,014 | 566 | 0.7 | 620 | 1.3 |
| Cote d'Ivoire | 1 | 1,359 | 901 | 1.2 | 617 | 1.3 |
| Croatia | 3 | 835 | 471 | 0.6 | 364 | 0.8 |
| Czech Republic | 3 | 811 | 419 | 0.5 | 395 | 0.8 |
| Dominican Republic | 2 | 3,194 | 1,326 | 1.7 | 1,870 | 4 |
| Ecuador | 2 | 2,357 | 1,572 | 2.1 | 939 | 2 |
| Estonia | 3 | 785 | 402 | 0.5 | 383 | 0.8 |
| Ethiopia | 1 | 2,099 | 1,817 | 2.4 | 307 | 0.7 |
| Georgia | 1 | 1,226 | 854 | 1.1 | 423 | 0.9 |
| Ghana | 1 | 2,466 | 1,614 | 2.1 | 881 | 1.9 |
| Hungary | 3 | 1,069 | 459 | 0.6 | 610 | 1.3 |
| India | 1 | 7,239 | 5,123 | 6.7 | 2,300 | 4.9 |
| Kazakhstan | 2 | 3,248 | 2,375 | 3.1 | 892 | 1.9 |
| Kenya | 1 | 3,168 | 2,237 | 2.9 | 962 | 2 |
| Lao People's Dem. R. | 1 | 1,683 | 745 | 1 | 963 | 2 |
| Latvia | 3 | 639 | 289 | 0.4 | 353 | 0.7 |
| Malawi | 1 | 3,855 | 2,448 | 3.2 | 1,416 | 3 |
| Malaysia | 3 | 3,715 | 1,968 | 2.6 | 1,765 | 3.7 |
| Mali | 1 | 446 | 196 | 0.3 | 302 | 0.6 |
| Mauritania | 1 | 1,452 | 691 | 0.9 | 936 | 2 |
| Mauritius | 3/ | 3,173 | 1,712 | 2.2 | 1,466 | 3.1 |
| Morocco | 2 | 2,886 | 1,959 | 2.6 | 927 | 2 |
| Myanmar | 1 | 2,089 | 1,667 | 2.2 | 422 | 0.9 |
| Namibia | 2 | 1,806 | 693 | 0.9 | 1,169 | 2.5 |
| Nepal | 1 | 4,798 | 3,336 | 4.4 | 1,480 | 3.1 |
| Pakistan | 1 | 4,786 | 3,686 | 4.8 | 1,159 | 2.5 |
| Paraguay | 2 | 3,823 | 2,426 | 3.2 | 1,398 | 3 |
| Philippines | 2 | 4,664 | 3,531 | 4.6 | 2,188 | 4.6 |
| Russian Federation | 2 | 3,148 | 1,844 | 2.4 | 1,331 | 2.8 |

Table 6.1 Country samples (alphabetic): total, outpatients (12 months), inpatients (5 years) and share of total dataset for each (continued)

| Country name | Country Level of Income* | Sample | Use of outpatient services | Percent of pooled data (outpatients) | Use of inpatient services | Percent of pooled data (inpatients) |
|----------------------|--------------------------|---------|----------------------------|--------------------------------------|---------------------------|-------------------------------------|
| Senegal | 1 | 803 | 479 | 0.6 | 470 | 1 |
| Slovakia | 3 | 1,515 | 1,045 | 1.4 | 508 | 1.1 |
| South Africa | 2 | 1,057 | 472 | 0.6 | 696 | 1.5 |
| Spain | 4 | 5,035 | 2,954 | 3.9 | 2,123 | 4.5 |
| Sri Lanka | 2 | 4,572 | 2,454 | 3.2 | 2,502 | 5.3 |
| Swaziland | 2 | 400 | 138 | 0.2 | 338 | 0.7 |
| Tunisia | 2 | 3,865 | 2,800 | 3.7 | 1,549 | 3.3 |
| Ukraine | 1 | 1,669 | 931 | 1.2 | 930 | 2 |
| United Arab Emirates | 4 | 753 | 481 | 0.6 | 291 | 0.6 |
| Uruguay | 3 | 1,741 | 1,055 | 1.4 | 739 | 1.6 |
| Viet Nam | 1 | 2,370 | 1,723 | 2.2 | 873 | 1.9 |
| Zambia | 1 | 3,102 | 2,217 | 2.9 | 894 | 1.9 |
| Zimbabwe | 1 | 2,596 | 1,773 | 2.3 | 848 | 1.8 |
| Total | | 119,162 | 76,634 | 100 | 47,092 | 100 |

* World Development Report 2003: 1= Low income; 2= Lower-middle income; 3= Upper-middle income; 4= High income

| Responsiveness Domain label (alternative name and short description) | Questions (response scale: very good, good, moderate, bad, very bad) "For your last visit,..." |
|---|--|
| Prompt attention (Convenient travel and short waiting times) | ...how would you rate the amount of time you waited before being attended to? |
| Dignity (Respectful treatment and communication) | ...how would you rate your experience of being greeted and talked to respectfully? |
| Communication (Clarity of communication) | ...how would you rate the experience of how clearly health care providers explained things to you? |
| Autonomy (Involvement in decisions) | ...how would you rate your experience of being involved in making decisions about your health care or treatment? |
| Confidentiality (Confidentiality of personal information) | ...how would you rate the way your personal information was kept confidential? |
| Choice (Choice including for continuity of health care provider) | ...how would you rate the freedom you had to choose your [health care provider]? |
| Quality of basic amenities (Basic amenities, quality of surroundings) | ...how would you rate the cleanliness of the rooms inside the facility, including toilets? |
| Access to family and community support (Social support, contact with outside world and maintenance of regular activities) | ...how would you rate the ease of having family and friends visit you? |

Figure 6.1 Responsiveness experience questions from the World Health Survey questionnaire (<http://www.who.int/healthinfo/survey/en/>)

Independent variables (covariates)

The 20 individual-level independent variables were also extracted from the WHS dataset: 12 questions for health service qualities, and 8 questions for personal characteristics. Health service variables are:

1. service management (government (1); private for-profit or other);
2. reason for admission (inpatient only: fever, severe diarrhoea, or cough [base category]; childbirth; arthritis, asthma, heart disease; bodily injury or minor surgery; other);
3. length of hospital stay (>5 days (1); 1-5 days);
4. perceived adequate quality of equipment (inpatient questionnaire) or adequate quality of medicines (outpatient questionnaire) (yes (1));
5. health care provider (outpatient only: doctor (1); nurse, dentist, or other);
6. sex of provider (outpatient only: female(1); male);
7. wait for admission (inpatient only: more than 1 week (1); less than 1 week);
8. number of people sharing a hospital room (inpatient only: continuous);
9. perceived discrimination because of age or sex or race or religion (yes (1));
10. mode of transport (public transport, bicycle or walking (1); private motor vehicle);
11. natural log of the time travelled (in minutes) to the health care facility;
12. timing of last visit (inpatient only: more than 1 year ago (1); less than 1 year ago).

Users' personal characteristics commonly included in previous studies were selected for the regressions to ensure consistency^{19,27}: age (≥ 30 yrs (1)); sex of respondent (male (1)); education (more than primary schooling completed (1)); marital status (single or not co-habiting (1)); caring for close friend or relative with chronic illness (1); household income (cross-country comparable permanent household income or asset index calculated by WHO with published descriptions elsewhere^{26,29} ranging from -4 to +3); and household location (rural (1)). SAH was dichotomized to reduce model estimation time for the larger outpatient dataset ("very bad", "bad", "moderate" =1) but retained as continuous variable (1 to 5 ordinal scale) for the inpatient dataset, noting previous research showing similar regression results for dichotomized or non-dichotomized SAH.²⁸

Data cleaning

Inpatient and outpatient visits were compiled into two different datasets for analysis (questionnaires required respondents to answer one or the other although there were some cases where interviewees unaccountably answered both sections (4% of respondents) and these cases were included in both the respective datasets as applicable). Missing data substitution procedures were applied to 5% of respondent records that contained more than 1 missing question (*mi* regression command in Stata (version 12)). Variables used in the estimation procedure were:

country; respondent sex; and place of residence (urban, rural). From 10 imputations, one series of imputed values was randomly extracted. For inpatient and outpatient services analyses, 46,477 and 76,426 final complete records were available for analysis.

Descriptive analyses

Responsiveness performance indicators, defined by measures of average levels and inequalities, were calculated for each domain. Average performance levels are simple frequencies. Inequality is calculated as relative gap measure, by domain, as is commonly used for SAH. Relative gap measures use country-specific income quintiles as elsewhere³⁰ with scores of the lowest income quintile (quintile 1) divided by scores for the highest income quintile (quintile 5).

Regression analysis

A mixed multi-level probit regression model was estimated for each of the 15 responsiveness domain measures (8 for inpatient visits, 7 for outpatient visits). Regression results reported are the independent variable coefficients and their statistical significance category (p -value <0.05). Positive coefficients indicate more 'problems' or worse responsiveness performance. The elaboration of results uses a pragmatic qualitative approach to describe findings across domains: if variable coefficients in roughly over half the domain regressions – 5 (4) out of 8 (7) inpatient (outpatient) – have p -statistics less than p -value =0.05, these coefficients are highlighted as stronger explanatory variables for responsiveness. To illustrate the sensitivity of a single domain as an example, the quality of basic amenities (outpatient) domain measure is plotted on a bar graph for categories within two strong explanatory (independent) variables.

Condition numbers are reported for the regression models. Condition numbers are defined as "the square root of the ratio of the largest to smallest eigenvalues of the Hessian matrix" (page 20, Rabe-Hesketh, Skrondal, Pickles 2004³¹) with large being as ≥ 50 -60 and small ≤ 15 -20. Statisticians note that large condition numbers do not necessarily imply poor model fit but that there are few instances of models with small condition numbers having poor fit.³¹ Log likelihoods (less negative, larger, implies better fit) are listed for each regression (all likelihood ratio chi-squares had p -values ≤ 0.000 , a test describing that at least one of the predictors are not zero). Model fit was also tested illustratively using predicted values from the quality of basic amenities domain.

The probit was incorporated into multilevel regression models as done in another published study.²⁰ The probit model assumes fixed scale cut-points across individuals (fixed on the single 'problem' score cut-point). Generally, the multilevel probit regression model defines two different estimates of error variances: individual (level 1) and country (level 2), setting the level 1 error (residual) variance equal to 1. In level 1, the explanatory variable coefficients are estimated.

In level 2, the country dummies are used to estimate the extent of additional error variance dependant on the respondent's country. This country-level coefficient is also known as the variance partitioning coefficient, ' ρ ', and commonly ranges between 0 and 0.3 (ρ between 0.0-0.10 implies a small effect).

Statistical software

Stata (version 12) was used for all analyses. The statistical software application used for the regression analyses was the Generalized Linear Latent and Mixed Models (GLLMM) software in Stata.³¹

RESULTS

Responsiveness descriptive statistics

Average levels of responsiveness problems are similarly ranked across domains for inpatient and outpatient services (Table 2, columns 1 (inpatient), 7(outpatient)). Inpatient services perform slightly worse in all domains, except in prompt attention. The worst performing domains with respect to the average levels are: autonomy (41% (inpatient), 39% (outpatient)); choice (45% (inpatient), 39% (outpatient)); and prompt attention (38% (inpatient); 43% (outpatient)). The best performing domains are dignity (24% (inpatient), 22% (outpatient)) and communication (26% (inpatient), 25%(outpatient)). Additional analyses comparing low income countries with other countries (see Appendix 6.1) show similar results, except for worse performance in inpatient confidentiality. The largest difference between inpatient and outpatient services is 6% for the choice domain.

Table 6.2 (columns 3, 9) shows the equity sensitivity of the responsiveness measures with users in the lowest income quintile having higher problems in all domains. Inequalities have similar ranges for inpatient and outpatient visits but slightly different domain rankings: the inpatients most equal domain score was 1.3 (basic amenities) and the least equal domain was 1.7 (dignity and confidentiality); the outpatients most equal domain score was 1.3 (basic amenities) and the least equal was to 1.8 (communication).

Regression results

Table 6.3(a) (inpatients) and 6.3(b) (outpatients) present the coefficients from the regression analyses. Level 1 explanatory variable coefficients (in rows) are listed for each regression analyses (columns). Coefficients are listed in three tiers: health-care service characteristics, the user's personal characteristics, and country-level effects (level 2, unexplained error variance).

Service characteristics versus users' personal characteristics

Worse inpatient responsiveness (Table 6.3(a): coefficients in 5 or more out of 8 domain regressions have p-statistics < 0.05) is explained by the following service characteristics (ordering follows table): management by the government; reason for admission, "fever, severe diarrhoea or cough"; perceived inadequate equipment; greater numbers of patients sharing rooms (space); and perceived discrimination. Users' personal characteristics explaining worse inpatient responsiveness are: age (youth); worse SAH; and lower income.

As seen in Table 6.3(b), worse outpatient responsiveness (coefficients in 4 out of 7 domain regressions have p-statistics < 0.05) is explained by the following service characteristics: management by the government; inadequate quality of medicines; perceived discrimination; and travel time to the facility. Personal characteristics associated with worse outpatient responsiveness are: age (youth); sex (male); worse SAH and lower income.

Overall across the 15 regressions, more health-care service than personal characteristics explain responsiveness scores (taking into consideration the differing number of variables in each tier 10 (inpatient) or 7 (outpatient) service characteristics variables versus 8 personal variables). Coefficient size is generally larger for health services characteristics than for the users' personal characteristics (most scales are comparable – see column headed "scale"). Inpatient service coefficients (p-value <0.05) range from 0.06 to 0.62 (absolute values) compared with 0.05 to 0.15 for variables describing users' personal characteristics.

Figure 6.2 shows two individual-level characteristics for which there are very similar patterns across the inpatient and outpatient regression results. The average level of outpatient basic amenities problems, used as an example, is higher in government-managed services, but income-based inequalities are lower in government-managed services.

Overall individual-level characteristics with diverging effects between inpatient and outpatient regressions are as follows: longer travel time to facilities and the users' sex (males with more problems) generally explain increased problems with outpatient rather than inpatient responsiveness.

Table 6.2 Responsiveness performance: average levels, inequalities and domain rankings (pooled data)

| Domains | Inpatient services (n=46,476) | | | | | | Outpatient/ambulatory services (n=76,426) | | | | | |
|------------------|--|---|---|--|--|------------------------------------|---|---|---|--|--|------------------------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | Frequency of users reporting a problem | Rank of worse performing domain (1=worst) | Ratio of inequality (users reporting a problem in quintile 1 divided by quintile 5) | Rank of worse performing relative inequality in domain (1=worst) | Ratio of inequality for non-government | Ratio of inequality for government | Frequency of users reporting a problem | Rank of worse performing domain (1=worst) | Ratio of inequality (users reporting a problem in quintile 1 divided by quintile 5) | Rank of worse performing relative inequality in domain (1=worst) | Ratio of inequality for non-government | Ratio of inequality for government |
| Prompt attention | 0.38 | 3.0 | 1.4 | 6.5 | 1.9 | 1.2 | 0.43 | 1.0 | 1.3 | 7.0 | 1.4 | 1.2 |
| Dignity | 0.24 | 8.0 | 1.7 | 1.5 | 2.2 | 1.3 | 0.22 | 7.0 | 1.7 | 2.0 | 2.3 | 1.3 |
| Communication | 0.28 | 7.0 | 1.6 | 3.5 | 2.4 | 1.3 | 0.26 | 6.0 | 1.8 | 1.0 | 2.4 | 1.3 |
| Autonomy | 0.41 | 2.0 | 1.6 | 3.5 | 2.1 | 1.3 | 0.39 | 2.5 | 1.5 | 4.3 | 1.8 | 1.1 |
| Confidentiality | 0.34 | 4.0 | 1.7 | 1.5 | 3.2 | 1.2 | 0.31 | 4.5 | 1.6 | 3.0 | 2.3 | 1.1 |
| Choice | 0.45 | 1.0 | 1.5 | 5.0 | 2.1 | 1.2 | 0.39 | 2.5 | 1.5 | 4.3 | 2.0 | 1.1 |
| Basic amenities | 0.34 | 5.0 | 1.3 | 8.0 | 2.3 | 0.9 | 0.31 | 4.5 | 1.5 | 4.3 | 2.2 | 1.0 |
| Social support | 0.30 | 6.0 | 1.4 | 6.5 | 2.2 | 1.1 | 0.33 | n/a | n/a | n/a | n/a | n/a |
| Average | 0.34 | | 1.5 | | 2.3 | 1.2 | 0.33 | | 1.5 | | 2.0 | 1.2 |

Table 6.3(a) 8 inpatient multilevel regressions^a explaining responsiveness problems (n=46,476): variable coefficients and country-level effect

| | | Scale | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice | Basic amenities | Social support |
|---|--|-----------|------------------|--------------|---------------|--------------|-----------------|--------------|-----------------|----------------|
| Independent variables (covariates)^b | | | | | | | | | | |
| Service characteristics | | | | | | | | | | |
| a1 | Recall period (=1 (<1 yr ago), 0 (more than 1 yr)) | 0/1 | 0.00 | 0.00 | 0.03 | -0.02 | -0.04 | -0.02 | -0.03 | 0.03 |
| a2 | Government (=1, versus 0, private for-profit/other) | 0/1 | 0.26 | 0.36 | 0.30 | 0.27 | 0.34 | 0.40 | 0.52 | 0.31 |
| a3 | Reason for admission [base: fever, diarrhea, cough] | | | | | | | | | |
| | -childbirth | 0/1 | -0.27 | -0.02 | -0.04 | -0.06 | -0.05 | 0.01 | -0.11 | 0.01 |
| | -arthritis, Asthma, Hrt dis | 0/1 | -0.10 | -0.15 | -0.16 | -0.07 | -0.14 | -0.12 | -0.12 | -0.14 |
| | -bodily injury & minor surgery | 0/1 | -0.10 | -0.01 | -0.17 | -0.16 | -0.12 | -0.16 | -0.11 | -0.10 |
| | -other | 0/1 | -0.06 | -0.02 | -0.03 | -0.03 | -0.03 | -0.06 | -0.11 | -0.06 |
| a4 | Length of hospital stay (=1 (>5 dys) versus 0, 1-5dys) | 0/1 | -0.02 | -0.02 | -0.07 | -0.02 | -0.02 | -0.04 | -0.10 | -0.03 |
| a5 | Adequate quality of equipment (=1) | 0/1 | -0.49 | -0.59 | -0.62 | -0.52 | -0.55 | -0.59 | -0.64 | -0.44 |
| a6 | Wait for admission (=1 (>wk) versus 0, less week) | 0/1 | 0.37 | 0.09 | 0.10 | 0.08 | 0.06 | 0.04 | 0.02 | 0.02 |
| a7 | Numbers sharing room | 1 - 50 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 |
| a8 | Perceived discrimination (=1) | 0/1 | 0.44 | 0.60 | 0.48 | 0.41 | 0.46 | 0.37 | 0.46 | 0.39 |
| a9 | Public transport, bicycle or walking (=1, vs. private motor vehicle) | 0/1 | 0.08 | 0.03 | 0.01 | 0.03 | 0.01 | 0.07 | 0.03 | 0.04 |
| a10 | Log of the travel time in minutes | -8.7- 4.5 | 0.07 | 0.02 | 0.01 | 0.01 | 0.00 | 0.01 | -0.01 | 0.05 |
| User's personal characteristics | | | | | | | | | | |
| 1 | Age, >30yrs (=1) | 0/1 | -0.12 | -0.09 | -0.09 | -0.08 | -0.08 | -0.13 | -0.05 | -0.12 |
| 2 | Sex of respondent, male (=1) | 0/1 | 0.00 | 0.03 | 0.00 | 0.00 | -0.01 | 0.09 | 0.00 | 0.02 |
| 3 | Education, more than primary schooling (=1) | 0/1 | -0.01 | 0.04 | 0.06 | -0.03 | -0.02 | 0.01 | 0.06 | 0.04 |
| 4 | Single or not co-habiting (=1) | 0/1 | -0.02 | -0.04 | 0.01 | 0.04 | -0.02 | -0.01 | 0.03 | -0.02 |
| 5 | Self-assessed health (higher worse=1) | 1-5 | 0.06 | 0.06 | 0.08 | 0.07 | 0.07 | 0.07 | 0.05 | 0.05 |

Table 6.3(a) 8 inpatient multilevel regressions^a explaining responsiveness problems (n=46,476): variable coefficients and country-level effect (continued)

| Independent variables (covariates) ^b | Scale | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice | Basic amenities | Social support |
|---|-----------|------------------|--------------|---------------|--------------|-----------------|--------------|-----------------|----------------|
| 6 Caring for close friend or relative with chronic illness (=1) | 0/1 | 0.01 | -0.03 | -0.06 | -0.03 | -0.02 | 0.01 | 0.05 | 0.04 |
| 7 Household income (assets factor score, low-high) | -4.0- 3.0 | -0.09 | -0.14 | -0.15 | -0.13 | -0.14 | -0.14 | -0.02 | -0.08 |
| 8 Household located in a rural area | 0/1 | 0.01 | -0.08 | -0.07 | 0.03 | -0.04 | 0.01 | -0.08 | -0.07 |
| Random effect / random intercept (mean error) | | 0.07 (0.02) | 0.06(0.01) | 0.07 (0.02) | 0.11(0.02) | 0.06(0.1) | 0.12(0.03) | 0.09(0.02) | 0.11(0.03) |
| Rho (variance partitioning coefficient; country dependency or country-level effect) | | 0.07 | 0.06 | 0.07 | 0.10 | 0.06 | 0.11 | 0.08 | 0.12 |

^a Model statistics are as follows (Domain: Condition Number, Log Likelihood); Prompt Attention: 60.171084, -27083.581; Dignity: 60.097013, -22083.746; Communication: 56.675267, -24428.14; Autonomy: 63.739473, -28388.151; Confidentiality: 57.22118, -26536.158; Choice: 65.75821, -29019.298; Quality of basic amenities: 62.7942, -25292.754; Social support: 65.75821, -29019.298; ^b Coefficients of covariates that have p-value<=0.05 are bolded and underlined

Table 6.3(b) 7 outpatient multilevel regressions^a explaining responsiveness problems (n=76,426): variable coefficients and country-level effect

| Independent variables (covariates) ^b | Scale | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice | Basic amenities |
|---|-----------|------------------|--------------|---------------|--------------|-----------------|--------------|-----------------|
| <i>Service characteristics</i> | | | | | | | | |
| b1 Government (=1, versus 0 private for-profit / other) | 0/1 | 0.34 | 0.35 | 0.27 | 0.29 | 0.26 | 0.44 | 0.46 |
| b2 Adequate quality of medicines (=1) | 0/1 | -0.42 | -0.41 | -0.43 | -0.45 | -0.37 | -0.43 | -0.35 |
| b3 Doctor (=1, versus 0, nurse, dentist, other) | 0/1 | 0.13 | -0.05 | -0.08 | -0.07 | -0.16 | -0.07 | -0.09 |
| b4 Sex of provider (=1 (female)) | 0/1 | 0.05 | 0.01 | 0.00 | 0.00 | 0.02 | 0.04 | 0.00 |
| b5 Perceived discrimination (=1) | 0/1 | 0.42 | 0.54 | 0.49 | 0.41 | 0.44 | 0.40 | 0.37 |
| b6 Public transport, bicycle or walking (=1) | 0/1 | 0.07 | -0.03 | 0.02 | 0.04 | 0.03 | 0.03 | 0.06 |
| b7 Log of the travel time in minutes | 8.7 - 4.5 | 0.12 | 0.05 | 0.03 | 0.02 | 0.03 | 0.02 | 0.02 |

Table 6.3(b) 7 outpatient multilevel regressions^a explaining responsiveness problems (n=76,426): variable coefficients and country-level effect (continued)

| Independent variables (covariates) ^b | | Scale | Prompt attention | Dignity | Communication | Autonomy | Confidentiality | Choice | Basic amenities |
|---|---|-----------|------------------|--------------|---------------|--------------|-----------------|--------------|-----------------|
| <i>User's personal characteristics</i> | | | | | | | | | |
| 1 | Age, >30yrs (=1) | 0/1 | -0.07 | -0.05 | -0.04 | -0.05 | -0.05 | -0.03 | -0.06 |
| 2 | Sex of respondent, male (=1) | 0/1 | 0.00 | 0.04 | 0.04 | 0.06 | 0.03 | 0.06 | 0.03 |
| 3 | Education, more than primary schooling (=1) | 0/1 | 0.00 | 0.00 | -0.04 | -0.06 | -0.04 | -0.04 | 0.03 |
| 4 | Single or not co-habiting (=1) | 0/1 | 0.01 | -0.01 | -0.03 | -0.03 | 0.02 | 0.00 | -0.03 |
| 5 | Self-assessed health (=1 (moderate, bad, very bad)) | 0/1 | 0.12 | 0.09 | 0.07 | 0.12 | 0.08 | 0.07 | 0.05 |
| 6 | Caring for close friend or relative with chronic illness (=1) | 0/1 | 0.00 | -0.04 | -0.10 | -0.06 | -0.09 | -0.06 | -0.05 |
| 7 | Household income (assets factor score, low-high) | 4.0 - 3.0 | -0.06 | -0.12 | -0.13 | -0.11 | -0.10 | -0.12 | -0.10 |
| 8 | Household located in a rural area | 0/1 | -0.05 | -0.07 | -0.05 | 0.01 | -0.03 | 0.00 | -0.09 |
| Random effect / random intercept (mean error) | | | 0.08 (0.02) | 0.10 (0.02) | 0.07 (0.02) | 0.10 (0.02) | 0.10 (0.02) | 0.12 (0.03) | 0.15 (0.02) |
| Rho (variance partitioning coefficient; dependency or country-level effect) | | | 0.07 | 0.09 | 0.07 | 0.09 | 0.09 | 0.06 | 0.17 |

^a Model statistics are as follows (Domain: Condition Number, Log Likelihood); Prompt attention: 11.365154, -47600.78; Dignity: 11.285483, -35163.442; Communication: 10.289886, -38752.411; Autonomy: 11.840901, -46695.139; Confidentiality: 11.832699, -41030.981; Choice: 12.671096, -46623.226; Quality of basic amenities: 13.004873, -40691.707; ^b Coefficients of covariates that have p-value<=0.05 are bolded and underlined

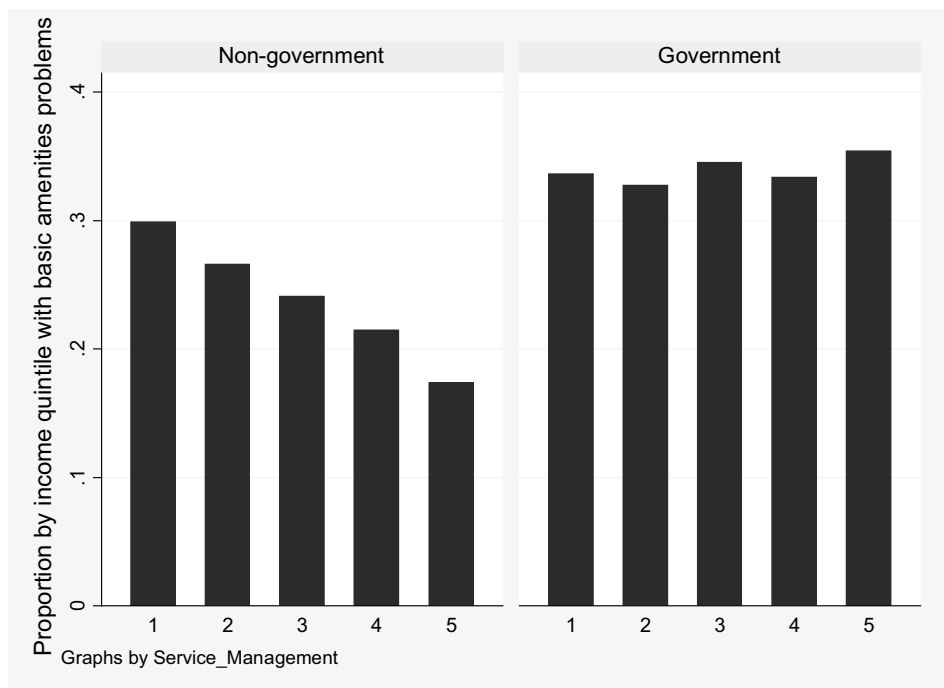


Figure 6.2 Average levels of responsiveness problems by country-specific income quintiles, stratified by management of health services: outpatient quality of basic amenities (n=74,282^a)

^a Note that a few observations were dropped from the final dataset (n=76,426) due to the missing data for the country-specific income variables in the WHO WHS dataset. This variable was not imputed.

Specific domain effects (*domain names italicised to facilitate reading*)

Inpatient setting

Government (versus non-government) management of hospitals is associated with worse choice and quality of *basic amenities*; childbirth on admission, with better *prompt attention*; waiting for admission for longer than one week, with worse *prompt attention*; sharing a hospital room with worse *basic amenities*; perceived discrimination with worse *dignity*. Being male is associated with worse choice; and living in a household in an *urban area* is associated with worse *dignity* and worse *basic amenities*.

Outpatient setting

Seeing a doctor (versus nurse) is associated with worse *prompt attention* but better *confidentiality*. Government management of services is associated with worse *choice* and *basic amenities*. Perceived discrimination is associated with worse *dignity*. Longer travel time is associated with worse *prompt attention*. Being male is associated with worse *communication*, *autonomy*, *confidentiality* and *choice*; being single and more educated is associated with better *autonomy*.

Country-level effects

Average country-level effects described in the last three rows of Table 3(a) and (b) are smaller for inpatient *prompt attention*, *dignity*, *communication*, *confidentiality* and *quality of basic amenities* domains. They are larger, overall, for outpatient services. The largest country-level effect is observed in outpatient *quality of basic amenities* where the variance partitioning coefficient, ρ , is 0.2.

DISCUSSION

This study analyses health system responsiveness to assess its association with health-care service and personal characteristics. The primary aim is to explore important service characteristics after taking full account of personal characteristics of users. The analysis covers all responsiveness domains in inpatient and outpatient settings and 12 service characteristics, as the key explanatory variables, for a dataset that covered almost 50 countries.

Service characteristics have pronounced associations with users' responsiveness domain scores at the individual-level controlling for the effects of users' personal characteristics. This finding implies that responsiveness domain measures, which use a common scale to describe non-clinical attributes of the patient experience, are capturing how service qualities affect the patient experience. Services characteristics studied here are commonly impacted by policy changes in different countries. Health service characteristics with particularly consistent associations with responsiveness are: discriminatory treatment by staff, less space availability (numbers of patients sharing rooms), and longer travel distances and time to the health facility. These characteristics all form part of health planning, legislation and contractual agreements related to health service provision, training and organization.

Users' appreciation of the interconnectedness between responsiveness and clinical quality was also apparent, a point of view which researchers have long argued on the basis of qualitative studies.^{7,32} This study found that user assessments of clinical quality are correlated with responsiveness but also differentiated from it, given the number of other explanatory factors that were also associated with responsiveness.

Responsiveness scores are clearly equity-sensitive. Real performance in responsiveness deteriorates for services confronted with very ill patients, who frequently combine multiple conditions (co-morbidity or multi-morbidity) with patient complexity (more complicated social backgrounds). The observation has been described by other researchers³³, and fits with everyday clinical experience, yet the fact that the association with responsiveness emerges in this complex analysis adds to its relevance. In our view, in spite of the known link of increasing age with less critical views³⁴, the association of younger ages and worse responsiveness probably also reflects true responsiveness performance, as the dichotomization procedure largely controlled for the reporting bias shifts related to age. Also, response bias coefficients associated with age are typically much smaller than the coefficients we obtained for age as a variable explaining responsiveness performance.³⁴ Therefore overall we think that this result illustrated that health systems may be systematically less well understood and poorly navigated by users in the 18 to 30 year old age groups, in particular male users. Behavioural changes associated with risk factors for noncommunicable diseases develop in these years, making the acceptability of services to these groups a very important consideration in improving health service programmes and coverage. In this respect, the strong tie of responsiveness to health outcomes represents both a warning and an opportunity.

The fact that government-managed services show poorer average performance (although with lower inequalities), carefully controlling for the population serviced, was not totally expected. Although this association is described in other studies, other studies have not fully accounting for the effects of the user's personal characteristics. One study in the literature that also used the WHS data documented lower average levels of responsiveness performance for countries with higher government expenditures as a share of total expenditures.¹⁷ Another systematic literature review study¹² reported lower quality in government-managed services but their definition of quality did not include the respect for persons responsiveness domains (autonomy, communication, confidentiality, dignity). Lower average levels of responsiveness in government-managed health services can probably be explained by their more restrictive expenditure per capita as domain-specific associations were strongest for the "client orientation" domains, which are known to be more associated with financial inputs. Interestingly, inequities (income-related) were clearly lower for government-managed outpatient services. This may be owing to the effective use of standardized protocols for dealing with patients. Overall income-related inequalities found for responsiveness (dignity, communication, confidentiality of information, and choice) supports calls by other researchers for more health systems interventions research in these areas.^{35,36}

While not being an aim of the study, criterion validity of the overall responsiveness concept was confirmed. First, the association between responsiveness and human rights, as described in the literature, was represented through a consistent association of perceived discriminatory

treatment and poorer responsiveness, with the dignity domain. This was a content validity relationship and not a question proximity effect as the question on dignity (greeting and respectful treatment) was 10 questions apart from the question for discrimination, with different wording ("treated worse because of.."). Second, other plausible associations between service characteristics and specific domains were also confirmed, for example, the effect of travel time on outpatient prompt attention. The importance of time as a barrier to access is highlighted in the literature, in particular for equity and preventive health-care services.⁹

Average country-level effects after individual-level variables are taken into account indicate the extent to which country-level variables that are not explicit in the model (e.g., total health expenditure, culture) may still explain service quality. Country-level effects were smaller for inpatient services and larger for outpatient services, in particular in the basic amenities regression. The smaller effects for inpatient services may be related to the skewed emphasis on tertiary care (hospital) expenditure in many developing countries and the skewed population finally obtaining access to hospitals. The overall larger country effect for outpatient services indicates more country-driven quality variation in primary health care services. The largest country-driven effect was recorded in the outpatient basic amenities regression, which probably reflects the role of health expenditure levels of countries in determining the more material aspects of health system responsiveness.

Several study limitations were present. In our view none of the limitations affect the main findings on the clear linkages between health-care service characteristics and responsiveness performance. First, the data are from surveys that are more than 10 years old. It is possible that in the intervening decade the relative performance of domains described by the descriptive statistics would have changed. For this reason we did not devote a large portion of the paper to elaborating on the performance results *per se* and rather the relationship between service and users characteristics and responsiveness. For example, one structural feature of health systems that could have affected prompt attention and communication in intervening years is the integration of information technology. This development could have improved communication and access. Unfortunately, the service characteristics described in the WHS did not include questions on information technology. However, other key performance findings, like worse levels but fewer inequalities in government-managed health services (relative to non-government managed services), are unlikely to have altered.

Related to the survey precision, survey respondent were asked to identify whether the health facility used was: (1) operated by the government; (2) privately operated; (3) NGO or (4) other. This question may be considered too difficult for respondents to answer accurately. Still, similar questions were asked of respondents in various waves of the Demographic and Health Surveys,

which have already been analysed and published.³⁷ The World Health Survey question followed the line of categorization used by the Organisation for Economic Co-operation and Development (OECD).³⁸ Also, lower responsiveness of government-managed services was found in another published systematic review that included more recent studies.¹²

The selection of the response cut-off to create the binary measure of responsiveness problems can be defended. As the dichotomization approach is a simple way to deal with reporting bias, it was better to select a higher threshold for quality, therefore grouping “moderate” with “bad” and “very bad” to signify a problem rather than grouping “moderate” with “good” and “very good”. Based on other published research²⁰, we could not use the “good” or “bad” as thresholds, as patterns of reporting bias show scale contraction, whereby some groups only use the “good”, “moderate” and “bad” categories. Also, the “bad” and “very bad” categories commonly contain only 5% of responses due to the skewed nature of the response data.

The final limitation was the complexity of the modelling. In developing the multi-level model, we tried unsuccessfully to include specific individual-level and country-level characteristics at the same time, which would be the most desirable approach. This was not computationally feasible. In future, perhaps a short-list of variables drawn from this study can be combined with selected country-level variables drawn from other studies. Nonetheless, understanding model fit in multi-level models such as we used is tricky. One approach is to compare forecast and raw data. We compared the pattern of forecast and raw data, stratified for non-governmental (steep gradient) or governmental management (no gradient) and found a relatively good fit in the domain of outpatient quality of basic amenities (shown in Appendix 6.2), which was the domain with the highest country-level variation. We felt that taking the extreme case was sufficient to demonstrate that the multi-level nature of the model was operating satisfactorily across domains, given that the other model fit statistics were also reviewed.

The evidence from this study arguably justifies the inclusion of responsiveness in health care information systems and related survey data collection instruments as is already recommended in some policy arenas.³² Our study shows associations between responsiveness and SAH at the individual-level, complementing studies describing associations with country-level population health outcomes.³⁹ The responsiveness domains described here can be used to develop a standard set of measures that are valid across diverse settings and which are sensitive to changes in service characteristics. This responsiveness information will complement information on financial barriers and service coverage rates, and can be used to help navigate health policy reforms and progress towards UHC.

CONCLUSIONS

Better real-time information from the users themselves will enable policy-makers to address service short-comings. Most current health monitoring systems contain information pertaining to utilization rates and out-of-pocket expenditures, without any assessment of the quality of the patient experience. Yet it is the financial means, combined with the clinical services and the human interaction that influences the ability of patients to receive care and recover from illness.

Several health service characteristics, including those addressing the human resources for health (or the health workforce agenda), are commonly the focus of policy interventions to improve quality or reduce costs. These system changes either enhance or diminish the quality of care from the user's perspective. The effects of these changes on patient experiences are commonly assumed, sometimes measured, but seldom monitored. Comprehensive reports covering responsiveness will allow all stakeholders to provide a better account of progress towards UHC (SDG 3.8), as well as related targets for promoting non-discrimination and inclusivity (e.g., SDG 16.b).

REFERENCES

1. de Silva A. A framework for measuring responsiveness. Global programme on evidence for health policy discussion paper no. 32. Geneva: World Health Organization; 2000. Accessed on 2 April 2016 from <http://www.who.int/healthinfo/paper32.pdf>.
2. Evans D, Hsu J, Boerma T. Universal health coverage and universal access. *B World Health Organ.* 2013;91:546-A.
3. World Health Organization, World Bank. Monitoring progress towards universal health coverage: framework, measures and targets. Geneva: World Health Organization; 2014.
4. Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bull World Health Organ.* 2000;78(6):717-31.
5. World Health Organization. Health systems: improving performance. The world health report 2000. Geneva: World Health Organization; 2000.
6. Aday L, Andersen R. A framework for the study of access to medical care. *Health Serv Res.* 1974;9:208-20.
7. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
8. Berwick, D. A User's Manual For The IOM's 'Quality Chasm' Report. *Health Affair.* 2002; 21(3):80–90.
9. Goddard M, Smith P. Equity of access to health care services: theory and evidence from the UK. *Soc Sci Med.* 2001;53:1149-62. Goddard M, Smith P. Equity of access to health care services: theory and evidence from the UK. *Soc Sci Med.* 2001;53:1149-62.
10. Gostin L, Hodge J, Valentine N, Nygren-Krug H. The domains of health responsiveness: a human rights analysis. Health and human rights working paper no. 2. Geneva: World Health Organization; 2003. [online: http://apps.who.int/iris/bitstream/10665/73926/1/HHRWPS2_eng.pdf]
11. Berwick D, Nolan T, Whittington J. The Triple Aim: care, health, and cost. *Health Affair.* 2008;27(3):759–69.
12. Basu S, Andrews J, Kishore S, Panjabi R, Stuckler, D. Comparative performance of private and public healthcare systems in low- and middle-income countries: a systematic review. *PLoS Med.* 9(6): p. e1001244.
13. Onukwugha E, Saunders E, Daniel Mullins C, Pradel FG, Zuck M. A qualitative study to identify reasons for discharges against medical advice in the cardiovascular setting. *BMJ Open.* 2012;2:e000902. doi:10.1136/bmjopen-2012-000902.
14. Shiferaw S, Spigt M, Godefrooi M, Melkamu Y, Tekie M. Why do women prefer home births in Ethiopia? *BMC Preg Childbirth.* 2013;13:5.
15. Üstün TB, Chatterji S, Mechbal A, Murray CJL, WHS Collaborating Groups. The World Health Surveys. Chapter 58. In: In Murray CJL, Evans DB, editors. Health Systems Performance Assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003. p.797–808.
16. King G, Murray CJL, Salomon J, Tandon A. Enhancing the validity and cross-cultural comparability of measurement in survey research. *Am Polit Sci Rev.* 2004 (February);98(1):191-207.

17. Robone S, Rice N, Smith P. Health systems' responsiveness and its characteristics: a cross-country comparative analysis. *Health Serv Res.* 2011;46(6pt2): 2079-100. doi: 10.1111/j.1475-6773.2011.01291.x.
18. Valentine NB, Prasad A, Rice N, Ribone S, Chatterji S. Health systems responsiveness: a measure of the acceptability of health-care processes and systems from the user's perspective. In: Smith P, Mossialos E, Papanicolaos I, Leatherman S, editors. *Performance measurement for health system improvement: experiences, challenges and prospects.* Cambridge: Cambridge University Press; 2010. p. 138-86.
19. Rice N, Robone S, Smith P. Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness. *Eur J Health Econ.* 2011;12:141-62.
20. Valentine NB, Verdes-Tennant E, Bonsel GJ. Health systems' responsiveness and reporting behaviour: multilevel analysis of the influence of individual-level factors in 64 countries. *Soc Sci Med.* 2015;138:152-60.
21. Peltzer K. Patient experiences and health system responsiveness in South Africa. *BMC Health Serv Res.* 2009;9(1):1-12. doi: 10.1186/1472-6963-9-117.
22. Gouveia G, Luna C, Souza-Júnior PR, Szwarcwald CL. Health care users' satisfaction in Brazil. *Cad Saude Publica.* 2005;21:109-18.
23. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet.* 2007;370:851-8.
24. Valentine NB, Bonsel GJ, Murray CJL. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Qual Life Res.* 2007;16(7):1107-25.
25. Rice N, Robone S, Smith P. Vignettes and health systems responsiveness in cross-country comparative analyses. *J R Stat Soc A Stat.* 2012;175(2): 337-69.
26. Hosseinpoor AR, Williams JA, Itani L, Chatterji S. Socioeconomic inequality in domains of health: results from the World Health Surveys. *BMC Public Health.* 2012; 12(198). doi: 10.1186/1471-2458-12-912.
27. Malhotra C, Kyung Do Y. Socio-economic disparities in health system responsiveness in India. *Health Policy Plann.* 2013;28(2):197-205.
28. Orly M, Matthews S, Power C. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *Int J Epidemiol.* 2000;29(1):149-57.
29. Gakidou E, Oza S, Vidal F, Li AY, Lee DK, Sousa A. Improving child survival through environmental and nutritional interventions: the importance of targeting interventions toward the poor. *JAMA.* 2007;298:1876-87.
30. World Health Organization. *State of Inequality.* Geneva: World Health Organization, 2015.
31. Rabe-Hesketh S, Skrondal A, Pickles A. *GLAMM Manual.* University of California Berkeley Division of Biostatistics Working Paper Series, Berkely: University of California, 2004. Accessed on 20 May 2016 from: <http://www.bepress.com/ucbbiostat/paper160>.
32. Bitton, A, Ratcliffe HL, Veillard JH, Kress DH, Barkley S, Kimball M et al. Primary Health Care as a foundation for strengthening health systems in low- and middle-income countries. *J Intern Med.* 2017; 32(5):566-71.

Chapter 6. Health System Responsiveness by N Valentine

33. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for understanding health and health services. *Ann Fam Med*. 2009;7(4): 357–63. doi: 10.1370/afm.983.
34. Fiorentini G, Ragazzi G, Robone S. Are bad health and pain making us grumpy? An empirical evaluation of reporting heterogeneity in rating health system responsiveness. *Soc Sci Med*. 2015;144 (November 2015):48–58. doi: 10.1016/j.socscimed.2015.09.009.
35. Gulliford M. Modernizing concepts of access and equity. *Health Econ Policy L*. 2009;4(pt2): 223–30. Doi:10.1017/S1744133109004940.
36. Cacace M, Ettelt S, Brereton L, Pedersen J, Nolte E. How health systems make available information on service providers. Experience in seven countries. Santa Monica: Rand, 2011. Accessed on 2 April 2016 from: http://www.rand.org/content/dam/rand/pubs/technical_reports/2010/RAND_TR847.pdf
37. Montagu D, Yarney G, Visconti A, Harding A, Yoong J. Where do poor women in developing countries give birth? A multi-country analysis of demographic and health survey data. *PLoS One*. 2001;6(2). doi: 10.1371/journal.pone.0017155.
38. Boehm K, Schmid A, Goetze R, Landwehr C, Rothgang H. Five types of OECD healthcare systems: empirical results of a deductive classification. *Health Policy*. 2013;113(3):258–69. doi: 10.1016/j.healthpol.2013.09.003.
39. Valentine NB, Bonsel, GJ. Exploring models for the roles of health systems' responsiveness and social determinants in explaining universal health coverage and health outcomes. *Glob Health Action*. 2016(Mar);Suppl 9:e29329.doi: 10.3402/gha.v8.29329.

Appendix 6.1 Description of responsiveness domain problems stratified and ranked for country subgroups of low income (23 countries) and other (26 countries): average domain levels, inequalities and rankings ^a

| INPATIENT | | | | | | | | |
|---------------------------------|-------------------------|-------------------|----------------------------|-----------------------|------------------------------|-------------------|----------------------------------|---------------------------|
| | Prompt attention | Dignity | Communi- cation | Auto- nomy | Confiden- tiality | Choice | Basic amen- ities | Social support |
| Other countries (n=27,048) | | | | | | | | |
| Mean level | 0.33 | 0.20 | 0.24 | 0.39 | 0.30 | 0.46 | 0.28 | 0.26 |
| Relative inequality | 1.4 | 1.7 | 1.7 | 1.6 | 1.7 | 1.6 | 1.2 | 1.2 |
| Rank - mean level | <u>3.0</u> | 8.0 | 7.0 | <u>2.0</u> | 4.0 | <u>1.0</u> | 5.0 | 6.0 |
| Rank- relative inequality | 6.0 | <u>1.3</u> | <u>1.3</u> | 4.5 | <u>1.3</u> | 4.5 | 7.5 | 7.5 |
| Low income countries (n=19,428) | | | | | | | | |
| Mean level | 0.43 | 0.30 | 0.33 | 0.44 | 0.37 | 0.44 | 0.41 | 0.34 |
| Relative inequality | 1.4 | 1.7 | 1.5 | 1.5 | 1.7 | 1.5 | 1.4 | 1.6 |
| Rank - mean level | <u>3.5</u> | 7.0 | 6.5 | 5.0 | <u>1.5</u> | <u>1.5</u> | <u>3.5</u> | 6.5 |
| Rank- relative inequality | 7.5 | <u>1.5</u> | 3.3 | 3.3 | <u>1.5</u> | 3.3 | 7.5 | 6.0 |
| OUTPATIENT | | | | | | | | |
| | Prompt attention | Dignity | Communi- cation | Auto- nomy | Confiden- tiality | Choice | Basic amen- ities | |
| Other countries (n=36,982) | | | | | | | | |
| Mean level | 0.43 | 0.17 | 0.20 | 0.35 | 0.24 | 0.39 | 0.23 | |
| Relative inequality | 1.4 | 1.8 | 1.8 | 1.4 | 1.5 | 1.5 | 1.4 | |
| Rank of mean level | <u>1.0</u> | 7.0 | 6.0 | <u>3.0</u> | 4.0 | <u>2.0</u> | 5.0 | |
| Rank- relative inequality | 5.3 | <u>1.5</u> | <u>1.5</u> | 5.3 | 3.5 | 3.5 | 5.3 | |
| Low income countries (n=39,444) | | | | | | | | |
| Mean level | 0.43 | 0.25 | 0.29 | 0.41 | 0.36 | 0.39 | 0.36 | |
| Relative inequality | 1.2 | 1.6 | 1.7 | 1.4 | 1.5 | 1.5 | 1.5 | |
| Rank of mean level | <u>1.0</u> | 7.0 | 6.0 | <u>2.0</u> | 4.5 | 3.0 | 4.5 | |
| Rank- relative inequality | 7.0 | <u>2.0</u> | <u>1.0</u> | 6.0 | 3.3 | 3.3 | 3.3 | |

^a The ranks of the two-to-three worst performing domains (1 – worst performing) in average levels and inequalities are bolded and underlined

Appendix 6.2 Model fit described using predicted forecast probabilities versus raw data percentages of having a problem with the quality of basic amenities in outpatient services by income quintile and service

| Quintile (1, low; 5, high) | Non-government services (private for-profit and not-for- profit) | | | Government services | | |
|----------------------------------|---|--------------------------|---------------------------------------|--|--------------------------|---------------------------------------|
| | Mean forecast probability from model | Number of respondents | Actual percentage from raw data | Mean forecast probability from model | Number of respondents | Actual percentage from raw data |
| 1 | 0.29 | 5,105 | 0.30 | 0.35 | 8,821 | 0.34 |
| 2 | 0.26 | 5,629 | 0.27 | 0.34 | 8,929 | 0.33 |
| 3 | 0.24 | 6,105 | 0.24 | 0.34 | 8,834 | 0.35 |
| 4 | 0.22 | 6,904 | 0.21 | 0.33 | 8,491 | 0.33 |
| 5 | 0.19 | 8,193 | 0.17 | 0.32 | 7,271 | 0.35 |
| Average | 0.24 | 31,936 | 0.24 | 0.34 | 42,346 | 0.34 |

CHAPTER 7

Exploring models for the roles of health systems' responsiveness and social determinants in explaining universal health coverage and health outcomes



N.B. Valentine and G.J. Bonsel

Global Health Action. 2016; 9:e29329

ABSTRACT

Background. Intersectoral perspectives of health are present in the rhetoric of the Sustainable Development Goals. Yet its descriptions of systematic approaches for an intersectoral monitoring vision, joining determinants of health and barriers or facilitators to accessing health care services are lacking.

Objective. To explore models of associations between health outcomes and health service coverage, and health determinants and health systems responsiveness, and thereby to contribute to monitoring, analysis and assessment approaches informed by an intersectoral vision of health.

Design. The study is designed as a series of ecological, cross-country regression analyses, covering between 23 and 57 countries with dependent health variables concentrated on the years 2002-03. Countries cover a range of development contexts. Health outcome and health service coverage dependent variables were derived from WHO information sources. Predictor variables representing determinants are derived from the WHO and World Bank databases; variables used for health systems' responsiveness are derived from the WHO World Health Survey. Responsiveness is a measure of acceptability of health services to the population, complementing financial health protection.

Results. Health determinants indicators – access to improved drinking sources, accountability, and average years of schooling – were statistically significant in particular health outcome regressions. Statistically significant coefficients were more common for mortality rate regressions than for coverage rate regressions. Responsiveness was systematically associated with poorer health and health service coverage. With respect to levels of inequality in health, the indicator of responsiveness problems experienced by the unhealthy poor groups in the population was statistically significant for regressions on measles vaccination inequalities between rich and poor. For the broader determinants, the Gini mattered most for inequalities in child mortality; education mattered more for inequalities in births attended by skilled personnel.

Conclusions. This paper adds to the literature on comparative health systems research. National and international health monitoring frameworks need to incorporate indicators on trends in and impacts of other policy sectors on health. This will empower the health sector to carry out public health practices that promote health and health equity.

INTRODUCTION

In the first decade of the 21st century the World Health Organization played a leading role in harmonizing health systems' performance assessment approaches through the development of relevant conceptual frameworks.¹⁻⁴ These frameworks refer to five health system goals that are achieved through the intermediate goal of coverage of the population with needed health services. According to these frameworks, different combinations of health systems' functions such as stewardship, financing or service delivery can be evaluated based on how well they improve the intermediate and final goals. The final goals are improvements in: population health levels, population health equity, levels of health systems' responsiveness to the legitimate expectations of the population, responsiveness equity, and fairness in financial contributions. When considering these frameworks and the associated monitoring approaches they have generated in the context of the Sustainable Development Goals (SDGs)^{5,6}, one can make several observations concerning potential areas for improvement. We focus on two areas for their improvement for purposes of this paper.

A first area for improvement in these frameworks is to address the neglect of the critical role of determinants beyond the health sector on population health. From the leading nineteenth century German doctor, Rudolf Virchow, to the present-day discussions on sustainable development, there is general recognition that average levels of population health and health inequities arise from factors beyond health care and the health systems' direct control. This implies augmentation of the original WHO frameworks mentioned above to include causal pathways beyond service coverage. As Hippocrates observed, social and environmental factors affect health directly. Yet social and environmental factors may give rise to additional problems with access to health services, thus modifying or even augmenting their direct effects on population health.

In order to be comprehensive and efficient, health performance frameworks and associated monitoring should track trends in these broader determinants. This will allow the health sector to detect, understand, influence, anticipate, and possibly even alter health impacts of decisions in other sectors. The WHO Commission on Social Determinants of Health argued in 2008 that impacts of health determinants, in particular social determinants related to the distribution of power, money and resources, were even more important for addressing health equity.⁷

A second area for improvement relates to the development of measures of non-financial barriers to access to health services. We use the term 'non-financial' to distinguish a set of barriers that complement the financing of direct medical expenses. So-called non-financial barriers may have components related to indirect costs (e.g., food, fear of loss of income), but also include other barriers related to acceptability and access (e.g., treatment with dignity and non-discrimination).

Non-financial barriers to health service access are related to health determinants. For example, the lack of transport in rural areas may result in longer travel distances to health facilities and differential health service access for disadvantaged groups. At the same time lack of transport can affect access to work with direct impacts on health through reducing family time or the length of periods of breast-feeding.

Although non-financial outcomes of health systems were reflected in WHO's original frameworks by the concept of health systems' responsiveness, advances in routine application in measurement and monitoring have been slow. Responsiveness is the degree to which legitimate expectations of the population with respect to non-clinical aspects of health care or public health services were actually met.¹ It is measured through large representative general population household surveys, or targeted surveys among recent care users. The responsiveness domains are, in alphabetical order: autonomy, choice, communication, confidentiality, dignity, prompt attention, (quality of) basic amenities, and (access to family and community) social support. The work of Donabedian, Tanahashi and others suggest that responsiveness has a direct positive relationship with service coverage and the final target, health.⁸⁻¹¹

We therefore plead for a broader measurement and monitoring framework, incorporating responsiveness and determinants, to be applied to evaluating health systems performance. To investigate the case for this empirically, this paper describes the development of analytical models that use data on health systems responsiveness and indicators of social and environmental determinants of health for their association with key outcomes from the original WHO frameworks. These key outcomes relate to: average levels of population health and health equity; and the intermediate goals of health service coverage and service coverage equity. These outcomes of interest are important in light of the SDGs as several measures of average levels of health and universal health coverage (UHC) have been accepted as part of the SDGs monitoring framework.¹² In our paper responsiveness and determinants are evaluated in terms of their instrumental contribution to health and health service coverage.

This paper investigates the association between population health outcomes, UHC, and responsiveness, and the role of determinants. We explore regression models, variables, and country-level indicators for determinants and responsiveness using cross-sectional data for between 23 and 57 countries. We observe whether a small basket of theory-supported determinants indicators explain expected linkages at the ecological level to health and coverage outcomes.

METHODS

The approach was: 1) to define a hypothesis-driven set of variables representing health service coverage, health, health systems responsiveness, health systems financial protection, and broader societal factors referred to as health determinants, suitable to test relationships; 2) to select and link accessible datasets for testing; 3) to conduct multiple regression analyses to assess the hypothesized associations. The country set was confined to those listed in the 57 face-to-face complete surveys of the World Health Survey (WHS) (2002-03) (see Appendix 7.1), for which comparable health systems responsiveness information is available.

Model

The analytical model that underpins the variables and regression analyses of population health and service coverage in this paper is represented by Figure 7.1.

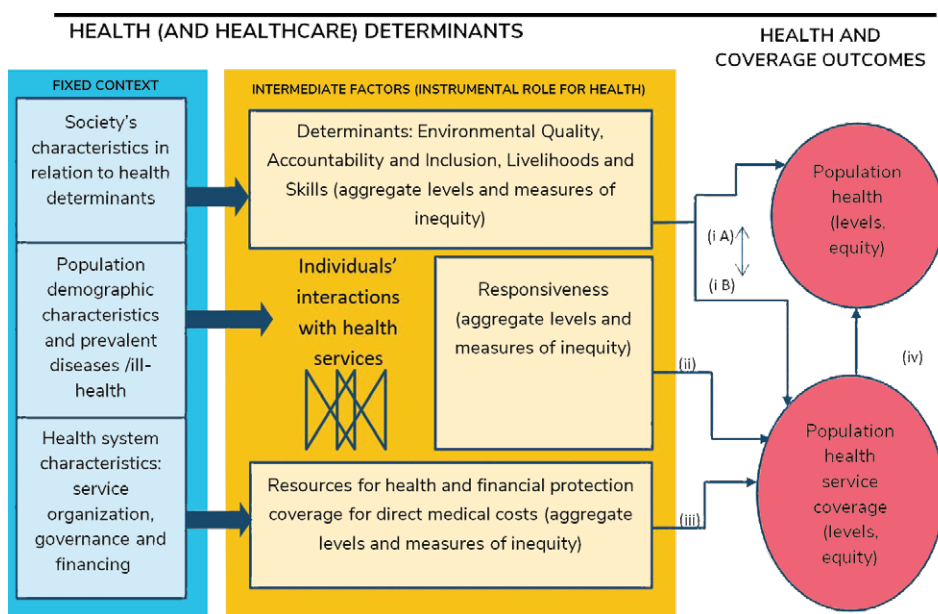


Figure 7.1 Analytical model for tracing key pathways of influence of determinants and responsiveness on population health and population health service coverage

Figure 7.1 describes the main pathways related to how determinants and responsiveness, as instrumental variables, affect population health and service coverage. It is derived from standard literature of conceptual models or frameworks related to how broader society interact with health systems to 'produce' health.^{7,13-15} In view of the broader analytical focus of this paper at this stage, these broader conceptual frameworks were considered to provide more relevant starting points than frameworks for monitoring of health services (e.g., monitoring of UHC). The societal level models tend to show the role of determinants quite strongly. The unique feature we added is the separate, instrumental and therefore testable role given to responsiveness, assuming that people-centeredness matters for health and coverage outcomes, thus representing a pathway itself.

From left to right, Figure 7.1 describes pathways of influence on population health and universal population health service coverage. Starting first with the blue block, the analytic framework assumes that 'fixed' characteristics of population, society, and the health systems' functions determine the context for intermediate factors that are more directly associated with health and coverage. These 'fixed' characteristics (left) are considered unchangeable for a given period of time, and result in multiple influences on intermediate factors. The intermediate factors shown in the centre column operate at the individual-level and include exposures or access to health services, for which empirical studies have shown more direct causative associations with population health (average levels and health equity) and health service coverage (average levels and coverage equity) (on the right).⁷ The intended analyses focus on the pathways: (i) determinants to health and coverage (not distinguishing between iA and iB); (ii) responsiveness to coverage; and (iii) financial resources and financial health protection to coverage. Pathway (iv) is assumed as implicit. Given the importance of financial protection for health service coverage, it was necessary to model, although the focus of the paper is on the additional roles of determinants and health systems responsiveness. Below we elaborate several generic implications of this analytical model for structuring the analyses that follow.

Health outcomes

1. Two outcomes should be considered in the regression models in order to cover two separate but important measures of health systems performance: population health and coverage of the population with essential health services (population health service coverage).
2. The outcome measures tested as the dependent variables should cover a spectrum of disease profiles and health service interventions.
3. Equity measures of these main coverage and health outcomes should also be considered in order to assess specific pathways for inequities in health and coverage.

Health determinants

1. Determinants as intermediate factors can be measured at the individual-level and aggregated to the country level, but they can also be measured by policy variables.
2. Distributional measures of determinants measured at the individual-level should be tested for modelling equity in health and health service coverage.
3. Health systems are key determinants of health and health service coverage. Proxy measures of health systems should be included in regression models and should relate to levels of resources and financial protection, given its importance as a determinant of health service clinical quality and access.⁵

The next sections focus on scoping recommended dependent and independent variables outlined in the analytical model and on selecting the final dataset for the regression models.

Scope of variables

Health and coverage (dependent variables)

Population health status can be represented by rates of morbidity, mortality, the compound indicator life expectancy or self-reported health. Health service coverage can be characterized by enrolment, utilization, or, effective service coverage rates (population in need receiving treatment divided by the population in need). Coverage rates can be measured comparably if morbidity-intervention combinations are standardized across countries.⁵ Similar to other studies¹⁶, the following additional criteria were applied to select the final set of dependent variables: 1) completeness of the data for the time period and countries; 2) a spectrum of health conditions or interventions covering reproductive, maternal and child health, communicable diseases (the so-called 'unfinished' Millennium Development Goals), and noncommunicable diseases and injuries; 3) variables for which country-level inequality data were available. Inequalities in health outcomes can be measured as gaps or concentration measures describing between-group differences in aggregate health outcome levels, where groups are defined by a 'wealth' or 'income' quintiles (e.g., absolute or relative gap between fifth and first quintiles, sex, geographic areas, educational attainment).¹⁷

Predictor and control variables (independent variables)

Health determinants variables (that are under the control of policy sectors other than the health sector) can be conveniently grouped into the following categories: I. Environmental Quality, II. Accountability and inclusion, and III. Livelihoods and skills (referred to as EQUAL). A variant of these categories was discussed at an expert meeting held by WHO.¹⁸ Based on these catego-

ries, there are a range of potentially relevant country-level indicators from a descriptive review of recent peer-review literature and from key informant reports.

Environment Quality indicators representing physical exposures are: urban households living in 'durable' structures; population exposed to small/fine urban particulates (PM10 or PM2.5) in concentrations exceeding WHO Air Quality Guidelines; households using modern fuels/technologies for all cooking, heating, and lighting activities; health facilities with access to clean and reliable electricity; population using a basic (improved) water source; the population whose access to safe water sources and sanitation is at risk from changing climate^{19,20}; exposure to harmful substances in the work environment; broader physical conditions in the work environment (e.g., night shifts, length of working week). Social elements of housing are residential stability, or affordability of neighbourhoods²⁰, urban design or green space and safety; and safety of products, enforceable and regulatory product quality and labelling measures.²¹

Accountability and inclusion, indicators include: violence against women; ratios of female to male schooling (attainment); social capital; self-reported gender inequality or discrimination; and discrimination in laws and policies, and related composite indices (e.g., World Bank Good Governance database).²²

Livelihoods and skills, indicators include: child stunting²³; caloric intake; household poverty; access to social protection (e.g., cash transfers); value in work and associated psychosocial exposures; employment relations (e.g., informal or formal, own account/salaried – access to paid parental leave, old age pensions); maternal education and birth spacing; child development; and access to early child development services, and social inequality.^{18, 23}

Responsiveness measurement is described in the literature.^{24,25} The original 8 responsiveness domains can be regrouped by the EQUAL framework: basic amenities and communication under Environmental Quality; autonomy, confidentiality, dignity, and social support under Accountability and inclusion; and choice and prompt attention under Livelihoods and skills.

Health systems pathways related to health system availability and financing are characterized in terms of levels of expenditure and financial protection coverage (other factors less commonly considered are human resource levels).^{4,16} Out-of-pocket expenditure indicators often represent financial protection coverage, with higher levels representing higher co-payments or low financial protection coverage^{16,26}, which are known to be regressive.²⁷

Demographic and biological drivers of need are primarily age and sex structure of the population. For our analysis such variables are controlled for as done elsewhere.²⁸

Data sources and final datasets

Country-level indicators and data

In view of the scope of variables and indicators outlined above, we scanned the range of potential data sources from WHO (World Health Statistics; Global Health Observatory) and World Bank (World Development Indicators (WDI), including the Worldwide Governance Indicators) databases. Country-level data were obtained, in most cases for the years 2002-03. For responsiveness, we needed to calculate country-level measures from individual-level datasets from the World Health Survey (WHS). The WHO WHS data is the only large publically available cross-country and region source with information on a range of health system responsiveness domains. Implemented between 2002 and 2004, the WHS data, acquired through nationally representative and quality-controlled surveys, have been widely used in the peer-review health literature.²⁹ Its data on responsiveness cover 57 countries and 151,848 respondents (using public and private sector providers). The selection of the remaining indicators were made for these 57 countries classified by the United Nations Development Agency in 2003³⁰: 23 low income countries; 13 lower middle-income countries; 11 upper middle-income countries; and 10 high income countries.

Table 7.1 lists the final indicator names, the number of observations obtained, descriptive statistics, and data sources.³¹⁻³⁴ All data except for responsiveness were obtained as country-level indicators. The estimation of country-level responsiveness indicators from the World Health Survey individual-level dataset³⁵ is described in detail below.

Acquiring and linking data from different sources took place between August and December 2014. Two consolidated datasets were used for analyses. The final 6 health and coverage average levels dataset contained between 52 (coverage) and 57 (health) country-level records. The final dataset for health and coverage inequalities consisted of 23 records (country-level).

Responsiveness indicators were derived from health service user responses to the WHS for all 57 countries as indicated earlier. Responsiveness level indicators were calculated by averaging domain summations of individual-level responses dichotomized from a 5-point verbal response scale ("very good", "good" [0, no problem]; "moderate", "bad", "very bad" [1, problem]). Dichotomizing the scale and standardizing by education and self-reported health status makes results less susceptible to 'reporting behaviour' bias and more comparable across countries.³⁶

Table 7.1 Variables used in regression models: descriptive statistics and data sources

| Analytic model categories | Variable or indicator names | Descriptive statistics | | | | Data source | |
|--|--|------------------------|----------|---------|---------|-------------|-----------|
| | | Mean | Std. Dev | Minimum | Maximum | Reference | Year |
| POPULATION HEALTH LEVELS | ALL (n=57) | | | | | | |
| | Maternal mortality per 100,000 live births (2005) | 308 | 368 | 1 | 1500 | (32) | 2005 |
| | Under five child mortality per 1000 live births (2005) | 63 | 66 | 4 | 220 | (33) | 2005 |
| | TB cause of death per 100,000 (2004) | 36 | 50 | 0.5 | 269 | (34) | 2004 |
| POPULATION HEALTH SERVICE COVERAGE LEVELS | ALL (n=52) | | | | | | |
| | Percentage of births attended by skilled health personnel (2000-06) | 76 | 28 | 6 | 100 | (32) | 2000-2006 |
| | Percentage of population covered with 1 dose of measles vaccination (2003) | 84 | 15 | 42 | 99 | (35) | 2003 |
| | Percentage of women receiving a pap smear (2000-06) | 31 | 29 | 0.1 | 82 | (32) | 2000-2006 |
| POPULATION HEALTH AND SERVICE COVERAGE EQUITY ^a | ALL (n=23) | | | | | | |
| | Child mortality: absolute difference by wealth quintile (poor quintile (I) less wealthy quintile (V)) | -57.6 | 32.8 | -157 | -15 | (32) | 1996-2006 |
| | Child mortality: relative ratio (wealthy quintile(I) / poor quintile (V)) | 0.5 | 0.2 | 0.3 | 0.8 | (32) | 1996-2006 |
| | Percent population with 1 dose measles vaccination: absolute difference by wealth quintile (wealthy quintile (I) less poor quintile (V)) | 24.7 | 13.6 | 1.9 | 46.9 | (32) | 1996-2006 |
| | Percent population with 1 dose measles vaccination: relative ratio (wealthy quintile / poor quintile) | 1.7 | 0.8 | 1 | 4.6 | (32) | 1996-2006 |
| | Percent live births with skilled personnel: absolute difference by wealth quintile (wealthy quintile (I) less poor quintile (V)) | 48.7 | 18.4 | 5.8 | 78.1 | (32) | 1996-2006 |
| | Percent live births with skilled personnel: relative ratio (wealthy quintile (I) / poor quintile (V)) | 6.3 | 8.2 | 1.1 | 38 | (32) | 1996-2006 |
| HEALTH AND HEALTH CARE DETERMINANTS – FIXED CONTEXT | FIXED CONTEXT (n=57) | | | | | | |
| | Accountability and voice (-2.5 to +2.5) | -0.07 | 0.96 | -2 | 1.6 | (31) | 2002 |
| | Control in limited regressions: Number of Lower income countries (2002) (n) | 23 | n/a | n/a | n/a | (31) | 2002 |
| | Control in limited regressions: Number of Lower middle income countries (2002) (n) | 13 | n/a | n/a | n/a | (31) | 2002 |
| | Control in limited regressions: Number of Upper middle-income countries (2002) (n) | 11 | n/a | n/a | n/a | (31) | 2002 |
| | Control in limited regressions: Number of High income countries (2002) (n) | 10 | n/a | n/a | n/a | (31) | 2002 |
| HEALTH AND HEALTH CARE DETERMINANTS – INTERMEDIATE | INTERMEDIATE FACTORS (n=57, except poverty) | | | | | | |
| | Access to improved drinking water (%) | 92 | 12 | 40 | 100 | (32) | 2000 |
| | Education (mean number of years) | 7.1 | 3 | 1 | 12.4 | (31) | 2000 |
| | Percentage of the population below the national poverty line (%) (n= 34) (2000-2006) | 37 | 15 | 6 | 69 | (31) | 2000-2006 |

Table 7.1 Variables used in regression models: descriptive statistics and data sources (continued)

| Analytic model categories | Variable or indicator names | | Descriptive statistics | | | | Data source | |
|---|--|---|------------------------|----------|---------|---------|-------------|-----------|
| | | | Mean | Std. Dev | Minimum | Maximum | Reference | Year |
| DETERMINANTS EQUITY MEASURES | Absolute difference in access to improved sources of drinking water (urban-rural) (n=23) | | 27 | 17 | -6 | 70 | (32) | 2000 |
| | Gini coefficient (0-1 index (1- highest income inequality) (n=23) | | 0.43 | 0.9 | 0.3 | 0.64 | (31) | 2000-2005 |
| RESPONSIVENESS | Aggregate level - percentage of responsiveness problems (%) | Dignity (n=57) | 22 | 11 | 6 | 53 | (36) | 2002/3 |
| | | Prompt Attention (n=57) | 35 | 12 | 16 | 67 | (36) | 2002/3 |
| | Inequality in responsiveness: difference by wealth or levels of responsiveness problems in poorest quintiles (IV, V) (outpatient services) (%) | Responsiveness level of problems in the poor quintile (I,II) (n=25) | 40 | 9 | 28 | 59 | (36) | 2002/3 |
| | | Absolute difference (Wealthy (V) less Poor (I,II)) (n=25) | 0 | 6 | 0 | 22 | (36) | 2002/3 |
| | | Relative ratio (Wealthy / Poor) (n=25) | 2 | 0 | 0 | 2 | (36) | 2002/3 |
| HEALTHCARE RESOURCES AND FINANCIAL PROTECTION (FOR MEDICAL COSTS) | Health expenditure per capita (International Dollars) (n=57) | | 624 | 837 | 21 | 3409 | (35) | 2002 |
| | Out-of-pocket health expenditure as a percentage of total health expenditure per cap. (n=25) | | 47 | 18 | 3 | 71 | (31) | 2002 |
| POPULATION DEMOGRAPHICS AND PREVALENT DISEASES | Population more than 60 years (%) (2006) (n=57) | | 11 | 7 | 2 | 24 | (32) | 2006 |

^a All wealth inequalities are based on household asset index quintiles (country-specific) calculated and provided by the data source listed

The final indicator calculated for the average level of responsiveness was: the frequency of reporting 'a problem' or 'poor responsiveness' in a particular domain. The domains of prompt attention and dignity were selected as they were among the two most important domains across a wide range of countries³⁷, and illustrated two different faces of responsiveness as described in the original WHO work (1): prompt attention, a "client orientation" domain, and dignity, a "respect for persons" domain. A composite responsiveness equity indicator was used for outpatient services rather than having domain-specific indicators. The responsiveness equity indicator was the average percentage across domains of responsiveness problems reported in the bottom two wealth quintiles for the less healthy in the population (those reporting moderate, poor, very poor health). Because of small numbers, the bottom two wealth quintiles were used rather than just the bottom). The wealth quintiles were based on cross-country comparable asset indices and made available by WHO as part of the World Health Survey dataset.³⁸ Like the poverty measure, this is not strictly speaking an inequality measure. But it does measure the responsiveness experiences of disadvantaged groups, which could explain inequities in health and coverage outcomes. Neither the relative gap or absolute gap measures of inequality for responsiveness showed any correlation with the dependent variables.

Missing data procedures

Missing data were not extensive for the final analyses. The missing data procedure followed used multiple imputation by chained equations as specified in the standard Stata *mi* command routines and associated instructions.³⁹ Missing data for the dependent (health, coverage) country-level indicators were not filled and the procedure was not necessary for the responsiveness and health systems indicators. Missing data for the determinants indicators were predicted from the country income group (dummy) and total health expenditure per capita. The following variables and observations were incomplete before imputation: accountability and voice index (missing for Cote d'Ivoire, Congo, Sri Lanka); and mean years of schooling for the population of 15 years or more in 2000 (which, of the variables filled, had the highest missing rates, for 9 out of 57 countries: United Arab Emirates, Bosnia and Herzegovina, Congo, Georgia, Israel, Lao People's Democratic Republic, Myanmar, Senegal, Tunisia). The percentage of the population below the national poverty line in 2002 was only available for 34 countries out of 57 countries and therefore not filled.

Analyses

Standard univariate and bivariate descriptive analyses on the dependent and independent variables preceded regression analysis (see Appendix 7.2). Normality of the distribution was tested. With respect to dependent variables, distributional characteristics required for several transformations. The logarithmic transformation of the dependent variables generally improved analytical properties. It was necessary to log the mortality rates in order to normalize the skewed

data distribution.^{16,40,41} For predictor variables, health expenditure per capita and the difference in access to improved water sources also required log transformation. These transformations do not affect the principle relationships tested. Scatterplots were also drawn to assess the linearity of bivariate associations between predictor and outcome variables. Correlations matrices were used to assess collinearity of predictors (the highest correlation was a Pearson correlation coefficient of 0.75 for access to improved water and log health expenditure per capita).

Different regression models were tested: ordinary least squares (OLS) linear regression, OLS log-linear regression, and Poisson and negative binomial maximum likelihood regressions. Although OLS regressions are more common than Poisson based models, it was appropriate to try different models based on assumptions regarding the outcome variable.⁴² Judging the appropriate form of the model of the outcome variables required assessing model fit statistics, considering the underlying data generation mechanisms assumption, as well as *a priori* assumptions regarding the impact of predictors on outcomes variables. Model comparisons were undertaken for the domains of dignity, prompt attention, and basic amenities as these variables have high importance and variance across countries.^{24,37} Model comparisons for mortality outcome variables, included log linear regressions and Poisson-based negative binomial models. The negative binomial is a form of the Poisson that recognizes the original count, and integer (non-negative) nature of data, while relaxing assumptions regarding the mean equal to the variance (high mean dispersion). It is arguably preferred for mortality regressions.^{42,43} Whereas OLS model fit statistic uses R-square, which ranges from 0 to 1, with numbers closer to 1 representing higher fit, the log-likelihood becomes more positive as fit improves. Comparing models using the log-likelihood statistic requires calculation of the likelihood ratio chi-square test (-2 times the difference in the log-likelihood ratios between the baseline and fitted models).

Only negative binomial regressions were used in regressions on average levels of health—maternal mortality, child mortality and Tuberculosis (TB) cause of death (mortality). For the aggregate levels of health coverage – population coverage of births by skilled attendants, coverage with measles vaccination, and receiving Pap smears – log-linear and linear regressions were used. Final regression models for health outcomes and overage levels contained a total of 6 predictor variables (after poverty rate was tested initially) and were each run twice in order to have separate predictions for dignity and prompt attention. This was done to reduce variables in a single regression, given the sample sizes of 57 and high correlations between responsiveness domain scores (Pearson correlation coefficient, 0.85).

For regressions on inequalities, final models had only 4 predictors at a time (only 23 countries). To select the 4 predictors, once again, the pre-testing of several models was performed. Regression results shown are selected from the model with the highest R-square or the most

positive log-likelihood ratios (best fit) from the 3 combinations of independent variables tested, which were: 1) out-of-pocket expenditure, responsiveness inequity, difference in access to drinking water between urban and rural areas, accountability and voice, Gini (largest number of variables); 2) out-of-pocket expenditure, responsiveness inequity, years of education (smallest number of variables); and 3) responsiveness inequity, the difference in access to drinking water between urban and rural areas, years of education, Gini. In all regressions a larger number meant greater inequity (favouring wealthier). In results, regressions were presented for difference and ratio properties of the 3 dependent variables (6 regression results).

Coefficients were assessed for statistical significance at the intervals: <0.10; <0.05; <0.001.

The negative binomial regression coefficients were interpreted as an increase of x in an explanatory variable multiplying the fitted mean mortality rate by $\exp(bx)$.⁴²

RESULTS

Comparing regression models for the role of predictors of health outcomes

Table 7.2 displays the regression test results using maternal mortality with the dignity domain for responsiveness as an example. Four regression formats are shown: ordinary least squares linear and log linear models (models 1,2 and 3), Poisson and negative binomial models (models 4,5 and 6). Comparisons of this nature were made for all outcome variables.

Using ordinary least square regression (regression 1), lower maternal mortality, without log transformation, is predicted by responsiveness but not by health expenditure per capita. This result is contrary to theory-driven expectations. In regression 2, with outcome variables log transformed, there is an association of maternal mortality with total health expenditure per capita (natural log), but the association for responsiveness is small and non-existent for years of schooling and access to drinking water.

Model 3, which also treats the outcome variable as logged, adds as an independent variable, the percentage of the population below the national poverty line, which was available for 34 mostly lower and lower middle income countries out of 57.

National poverty rates are associated with maternal mortality (coefficient 0.03 (p-value=0.00)) as would be expected. Again, the effect of national poverty rates swamps out health expenditure per capita and accountability, which may suggest that the log transformation alone is insufficient to correct for the underlying distributional form.

Comparing the model fit statistics for OLS regression shows better fit for log-linear regressions (model 2). The Poisson regression log-likelihood statistics indicate poor fit relative to the negative binomials (more negative). In the negative binomial regressions, compared with baseline models (containing only population over 60 years), both regression models likelihood ratio tests are adequate to warrant inclusion of more predictors (p-value <0.000).

Specific experimentation showed that for all models the percentage of the population older than 60 years is significantly associated with the level of maternal mortality. This obvious demographic-biological need pathway will receive no further comment. Other variables show less uniform patterns.

Regressions 4 and 5 introduces the country's average level of income (World Bank categories) as independent variables. Income removes some effects of other variables, in particular for health expenditure per capita, responsiveness and years of schooling. Regression 6, on the other hand, shows up these variables by dropping the income categories. Maternal mortality across countries is associated with health expenditure (coefficient: 0.26; p-value: 0.03) and with responsiveness barriers (coefficient: 0.29; p-value: 0.05) and average years of schooling (coefficients:-0.07; p-value: 0.08). While associations with accountability and access to water and sanitation are insignificant, the clearer pathways related to health systems and health service interactions, make this model preferable in our view, given that the health outcome maternal mortality is usually associated with the lack of health service attention at birth.

Table 7.2 Maternal mortality cross-country regression models using the responsiveness dignity domain only (percentage of problems reported by health service users)^a

| Regression no. | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | |
|--|-----------------------------------|--|---|--------------------------|---|---|---|
| Model | Ordinary least squares regression | Log-linear ordinary least squares regression | Log-linear ordinary least squares regression (with Poverty) | Basic Poisson regression | Negative binomial maximum likelihood regression | Negative binomial maximum likelihood regression | Comparing coefficient in models 2 and 6 |
| Fit | | | | | | | |
| MSE | 190.8 | 0.82 | 0.93 | | | | |
| R2 (or pseudo) | 0.76 | 0.86 | 0.5 | 0.93 | 0.17 | 0.14 | |
| Log Likelihood | | | | -6776 | -359 | -373 | |
| Average total health expenditure per capita (log) | | | | | | | |
| Coefficient | -19.54 | -0.4 | -0.03 | 0.18 | 0.02 | -0.26 | -0.40; -0.26 |
| Std. Error | 38.36 | 0.16 | 0.14 | 0.01 | 0.09 | 0.12 | |
| T-statistics | -0.51 | -2.42 | -0.21 | 35.14 | 0.22 | -2.17 | |
| P-value | 0.61 | <u>0.02</u> | 0.84 | <u>0</u> | 0.83 | <u>0.03</u> | |
| Percent population with responsiveness problems | | | | | | | |
| Coefficient | 110.39 | 0.33 | 0.34 | -0.17 | 0.31 | 0.29 | 0.33; 0.29 |
| Std. Error | 46.84 | 0.2 | 0.16 | 0.01 | 0.11 | 0.14 | |
| T-statistics | 2.36 | 1.63 | 2.09 | -32.28 | 2.75 | 1.97 | |
| P-value | <u>0.02</u> | 0.11 | <u>0.05</u> | <u>0</u> | 0.01 | <u>0.05</u> | |
| Percent population accessing drinking water | | | | | | | |
| Coefficient | -592.68 | -0.81 | -0.24 | -0.42 | -0.21 | -0.48 | -0.81; -0.48 |
| Std. Error | 133.13 | 0.57 | 0.44 | 0.01 | 0.32 | 0.46 | |
| T-statistics | -4.45 | -1.42 | -0.55 | -42.29 | -0.66 | -1.04 | |
| P-value | <u>0</u> | 0.16 | 0.59 | <u>0</u> | 0.51 | 0.3 | |
| Accountability and voice | | | | | | | |
| Coefficient | 74.11 | 0.01 | -0.09 | 0.25 | -0.06 | -0.04 | 0.01; -0.04 |
| Std. Error | 41.4 | 0.18 | 0.18 | 0 | 0.12 | 0.13 | |
| T-statistics | 1.79 | 0.04 | -0.51 | 92.95 | -0.52 | -0.3 | |
| P-value | 0.08 | 0.97 | 0.61 | <u>0</u> | 0.6 | 0.77 | |
| Average years of schooling | | | | | | | |
| Coefficient | -28.53 | -0.05 | -0.1 | -0.05 | -0.11 | -0.07 | -0.05; -0.07 |
| Std. Error | 12.85 | 0.06 | 0.05 | 0 | 0.03 | 0.04 | |
| T-statistics | -2.22 | -0.86 | -1.99 | -34.89 | -3.28 | -1.73 | |
| P-value | 0.03 | 0.4 | 0.06 | 0 | 0 | 0.08 | |

Table 7.2 Maternal mortality cross-country regression models using the responsiveness dignity domain only (percentage of problems reported by health service users)^a (continued)

| Regression no. | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | |
|---|-----------------------------------|--|---|--------------------------|---|---|---|
| Model | Ordinary least squares regression | Log-linear ordinary least squares regression | Log-linear ordinary least squares regression (with Poverty) | Basic Poisson regression | Negative binomial maximum likelihood regression | Negative binomial maximum likelihood regression | Comparing coefficient in models 2 and 6 |
| Percentage population over 60 years | | | | | | | |
| Coefficient | -13.51 | -0.15 | -0.14 | -0.25 | -0.17 | -0.18 | -0,15 -0,18 |
| Std. Error | 6 | 0.03 | 0.03 | 0 | 0.02 | 0.02 | |
| T-statistics | -2.25 | -5.84 | -4.65 | -151.75 | -10.01 | -8.67 | |
| P-value | 0.03 | 0 | 0 | 0 | 0 | 0 | |
| Percent population below national poverty line coefficient | | | | | | | |
| Coefficient | | | 0.03 | | | | |
| Std. Error | | | 0.01 | | | | |
| T-statistics | | | 3.4 | | | | |
| P-value | | | 0 | | | | |
| Income group (low-middle) | | | | | | | |
| Coefficient | | | | -0.91 | -0.95 | | |
| Std Error | | | | 0.01 | 0.25 | | |
| T-statistics | | | | -89.84 | -3.74 | | |
| P-value | | | | 0 | 0 | | |
| Income group (middle) | | | | | | | |
| Coefficient | | | | -1.73 | -1.97 | | |
| Std. Error | | | | 0.02 | 0.34 | | |
| T-statistics | | | | -107.24 | -5.83 | | |
| P-value | | | | 0 | 0 | | |
| Income group (high) | | | | | | | |
| Coefficient | | | | -2.71 | -2.77 | | |
| Std. Error | | | | 0.09 | 0.46 | | |
| T-statistics | | | | -30.57 | -6.05 | | |
| P-value | | | | 0 | 0 | | |
| Constant | | | | | | | |
| Coefficient | 3530.55 | 12.66 | 7.61 | -2.4 | -1.27 | -0.48 | |
| Std. Error | 515.92 | 2.21 | 1.91 | 0.03 | 1.23 | 1.78 | |
| T-statistics | 6.84 | 5.72 | 3.98 | -75.58 | -1.03 | -0.27 | |
| P-value | 0 | 0 | 0 | 0 | 0.31 | 0.79 | |

^a n is 57 countries for all regressions except for model 3, where the number of country observations is 34

Regression models which explain aggregate health levels

Six regressions of health outcomes and coverage rates are shown in Tables 7.3 and 7.4 for 2 domains: 1) dignity and 2) prompt attention. Regressions otherwise contain the same independent variables. The first regression column shows the maternal mortality negative binomial regression with dignity (repeated from Table 7.2, model 6). The next column in Table 7.3 shows the regression for child mortality, then TB mortality rates, and so on.

Across Table 7.3 and Table 7.4, one observes that statistically significant coefficients for predictors are more common for mortality rate regressions than for coverage rate regressions. In Tables 7.3 and 7.4, column 3, there are a high number of significant covariates, in the expected direction, for child mortality on the one hand (all predictors except adult education), and a low number for measles coverage (Tables 7.3 and 7.4, column 6). Responsiveness is statistically significant for all mortality regressions but for only one of the service coverage regressions (skilled attendants). Higher percentages of responsiveness problems in countries are associated with increased maternal, TB, and injuries mortality (Table 7.3, columns 2-4; Table 7.4, columns 2,4), and reduced coverage of the population with skilled birth attendants (Table 7.3, column 5). On average the effect sizes of responsiveness on the dependent variable, measured in terms of numerical percentages, are higher for service coverage than for mortality rates. Using coefficient results in model 6 of Table 7.2, and the proportionate formula for interpreting negative binomial regression coefficients (column 1 of Table 7.3), as described in the methodology [$\exp b(x) = \exp(0.29 \cdot 0.10)$], an increase in responsiveness problems by 10% increases (maternal) mortality rates by 3% and decreases service coverage rates by 5%.

Looking across indicators, access to improved drinking sources is statistically significant in only 1 mortality regression – child mortality and 1 coverage regression – skilled attendants at birth. Accountability (and voice) is statistically significant only in 1 mortality regression – child mortality. Average years of schooling is relevant to 1 mortality outcome – maternal mortality – and to 1 coverage outcome – measles vaccination. Health expenditure per capita is statistically significant in 4 out of 6 regressions (except TB mortality and measles coverage).

Model fit within health outcomes regressions, as judged by log-likelihood statistics is best for maternal mortality (LL= -359), followed by TB mortality (LL= -503) and child mortality. In health service coverage regressions, fit as judged by the R-square statistic, is better for skilled attendants at birth (R-square= 0.81) and Pap smear (R-square= 0.55) than measles vaccination (R-square= 0.37).

Qualitative changes are observed in the effects of dignity versus prompt attention. For child mortality, prompt attention barriers are not significant, whereas dignity barriers are significant

(p -value<0.05). On the other hand, both dignity and prompt attention barriers are highly significant (p -value<0.001) for TB mortality rate regressions. However, effect sizes for TB are larger for prompt attention responsiveness barriers than for dignity.

Table 7.3 Cross-country regression models for health outcomes and health service coverage, using the responsiveness domain dignity only ^a

| Explanatory variables | HEALTH OUTCOMES | | | SERVICE COVERAGE | | |
|---|---|--|---|--|--|----------------------------------|
| | Maternal Mortality (natural log-by negative binomial model) | Child Mortality (natural log-by negative binomial model) | TB Mortality (natural log-by negative binomial model) | Percentage coverage by skilled attendants at birth (natural log) | Percentage coverage of measles vaccination (natural log) | Percentage coverage of pap smear |
| Health expenditure per capita (log) | -0.26 ** | -0.21 ** | -0.2 | 0.48 ** | -0.15 | 9.24 ** |
| Users reporting responsiveness problems | 0.29 ** | 0.21 * | 0.74 *** | -0.49 ** | -0.25 | -6.79 |
| Access to improved drinking sources/water | -0.48 | -0.75 * | -0.69 | 2.53 *** | 1.22 | -8.06 |
| Accountability and voice (-2.5 to +2.5) higher better | -0.04 | -0.18 * | -0.19 | -0.32 | 0.05 | -0.97 |
| Average years of schooling of adults (>18yrs) | -0.07 * | -0.04 | 0.07 | 0.11 | 0.16 ** | 1.14 |
| Percentage of population over 60 years of age | -0.18 *** | -0.07 *** | -0.11 *** | 0.1 *** | 0.05 | 1.12 * |
| Constant | -0.48 | 2.56 * | -2.87 | -14.03 *** | -4.36 | -14.83 |
| Model fit | Negative binomial | Negative binomial | Negative binomial | Log-linear | Log-linear | Ordinary least squares |
| MSE | | | | 1.01 | 1.13 | 18.53 |
| R2 (pseudo for Poisson, NB) | 0.16 | 0.09 | 0.07 | 0.81 | 0.37 | 0.59 |
| Log Likelihood | -359 | -563 | -503 | | | |

^a n=57 except for regression for skilled birth attendants (n=52) * p -value<0.10; ** p -value<0.05; *** p -value<0.001

Table 7.4 Cross-country regression models for health outcomes and health service coverage, using the responsiveness domain prompt attention only^a

| Explanatory variables | HEALTH OUTCOMES | | | SERVICE COVERAGE | | |
|---|---|--|---|--|--|----------------------------------|
| | Maternal Mortality (natural log-by negative binomial model) | Child Mortality (natural log-by negative binomial model) | TB Mortality (natural log-by negative binomial model) | Percentage coverage by skilled attendants at birth (natural log) | Percentage coverage of measles vaccination (natural log) | Percentage coverage of pap smear |
| Health expenditure per capita (log) | -0.28 ** | -0.22 ** | -0.22 | 0.55 *** | -0.04 | 10.61 *** |
| Users reporting responsiveness problems | 0.31 * | 0.22 | 0.83 *** | -0.34 | 0.23 | -2.52 |
| Access to improved drinking sources/water | -0.5 | -0.76 ** | -0.53 | 2.6 *** | 1.22 | -7.28 |
| Accountability and voice (-2.5 to +2.5) higher better | -0.08 | -0.21 ** | -0.26 | -0.25 | 0.05 | -0.26 |
| Average years of schooling of adults (>18yrs) | -0.07 * | -0.04 | 0.06 | 0.1 | 0.15 ** | 1.03 |
| Percentage of population over 60 years of age | -0.18 *** | -0.07 *** | -0.11 *** | 0.1 *** | 0.05 | 1.1 * |
| Constant | -0.52 | 2.5 * | -3.71 | -14.14 *** | -4.48 | -16.81 |
| Model fit | Negative binomial | Negative binomial | Negative binomial | Log-linear | Log-linear | Ordinary least squares |
| MSE | | | | 1.03 | 1.13 | 18.89 |
| R2 (pseudo for Poisson, NB) | 0.16 | 0.09 | 0.07 | 0.8 | 0.37 | 0.57 |
| -2 times the Log Likelihood | -359 | -563 | -503 | | | |

^a n=57 except for regression for skilled birth attendants (n=52); *p-value<0.10; **p-value<0.05; ***p-value<0.001

Regression models which explain aggregate health inequalities

Table 7.5 shows regression results for health outcome inequalities and service coverage inequalities as dependent variables. Child mortality favoured combinations of variable sets 1 and 3, while coverage inequality models were best fitted with the smaller set of predictor variables from set 2.

Table 7.5 Cross-country regressions explaining inequalities in health status and health service coverage by contextual and instrumental factors including responsiveness (n=23)

| Explanatory variables | CHILD MORTALITY RATES | | BIRTHS ATTENDED BY SKILLED PERSONNEL | | MEASLES VACCINATION COVERAGE | |
|---|--|---------------------------------|--|----------------------------------|--------------------------------------|---------------------------------|
| | Difference ABS (poor-rich), larger worse | Ratio (poor/rich), larger worse | Diff-erence (rich-poor), larger, worse | Ratio (rich/poor), larger better | Difference (rich-poor), larger worse | Ratio (rich/poor), larger worse |
| Model | 3 | 1 | 1 | 2 | 2 | 2 |
| Out-of-pocket health expenditure | | 0.01 | 0.04 | 0.01 | 0.27 ** | 0.01 ** |
| Responsiveness problems (% unhealthy, poor) | 0.02 | -0.49 | 0.81 | 0.11 | 0.31 | 0.01 ** |
| Difference in percent population accessing drinking water | 0.01 | 0 | -0.1 | | | |
| Accountability index | | 0.50 (0.15) | 14.98 (0.12) | | | |
| Average years of schooling | 0.11 | | | -1.54 ** | -2.72 ** | -0.07** |
| GINI (0-1, 1 unequal) | 0.05 ** | 0.02 | -0.9 | - | - | |
| Model fit | Negative binomial | Ordinary least squares | Ordinary least squares | Ordinary least squares | Ordinary least squares | Log linear |
| MSE | | 0.74 | 20.43 | 7.46 | 9.53 | 0.24 |
| R2 (pseudo) | 0.05 | 0.11 | 0.03 | 0.15 | 0.5 | 0.55 |
| Log likelihood | -118.09 | | | | | |

*p-value<0.10; **p-value<0.05; ***p-value<0.001

Responsiveness problems experienced by the unhealthy poor groups was statistically significant only for measles vaccination inequalities between rich and poor. Out-of-pocket expenditure was statistically significant in predicting coverage gaps for measles immunization. With respect to the broader determinants, the Gini coefficient mattered most for inequalities in child mortality between the rich and poor, and education mattered more for inequalities in births attended by skilled personnel.

DISCUSSION

This paper presents an exploration of different models for understanding the linkages between health and service coverage outcomes, and related health determinants, including health systems responsiveness. We used a set of cross-sectional analyses of different types of health status and health service coverage rates to explore different sets of determinants and health systems responsiveness indicators across 57 countries. To our knowledge, this is the first time that both health conditions and service coverage rates are explained using determinants and 'acceptability' barriers of responsiveness.

The determinants' indicators tested here were associated with health in the expected directions as shown elsewhere^{17,20,21,40,41}, which is reassuring with respect to the findings for health systems responsiveness. An interesting new finding is that responsiveness was systematically associated with poorer health outcomes and coverage in the areas of maternal mortality, child mortality, TB mortality, skilled birth attendance coverage, and Pap smears (not measles vaccination). The results imply that both responsiveness barriers and health determinants have quantifiable, separate associations with health status and health service coverage. Responsiveness complements the financial barriers indicators recommended to be measured as part of UHC in the health goal, SDG-Goal 3.

Our analyses also have implications for monitoring health determinants in the SDGs. SDG-Goal 3 (health) covers health outcomes and 'UHC'.^{6,12} UHC in the SDGs is defined as the degree to which health services meet population health care needs without undue financial hardship. Two metrics derived for its quantification are: financial protection coverage of individuals, which is measured by the absence of so-called catastrophic direct medical costs⁵, and service or intervention coverage, which is measured as the proportion of people, who need particular well-accepted health interventions, receiving them. Both metrics can also be expressed as coverage inequality (by sex, education, income, geographic).¹⁷ Yet these metrics do not explicitly track responsiveness barriers, nor the wider panorama of social and environmental determinants such as education of mothers, income inequality, which are clearly important for achieving good population health and effective health service coverage.

The systematic testing of regression models, variables and indicators, as illustrated in this paper, is useful for determining which national comparable health determinants indicators to track. Our findings show that several determinant indicators are candidates. These include drinking water coverage and coverage inequalities, poverty, mean years of schooling and income inequality. These are candidates for both international and national use in intersectoral monitoring frameworks that track health determinants. Except for poverty, the data series are relatively

complete (poverty is more complete now relative to 2002-04) and they complement SDG - Goal 3 (health). These indicators are also likely to be used by sectors beyond health in monitoring other SDG goals (e.g., Goal 1- poverty, Goal 4-education, Goal 6-water and sanitation and Goal 10-income inequality). Having the health sector in national contexts track a set determinants indicators is vital, as described in the Health in All Policies approach.⁴⁴ Tracking determinants is statistically simple, as well as efficient and provides a rationale for policy coherence if the same indicators are already being used by another sector to monitor their strategic performance. These data can also be used as a bridge to build better information systems for health impact assessments, thereby enabling anticipation of health changes before they emerge as behavioural changes in the population.

There are several limitations to our study. It consists of data that are 12-13 years old. It is possible that in this time, as health systems and development contexts have changed, other patterns would emerge if the study were conducted on current data (e.g., governance accountability concepts can have altered). We also used a limited number of variables. More recommendations for the use of variables in future research is discussed below. A further limitation is that we only conducted relatively simple cross-sectional analyses, which yielded information on associations but specific longitudinal analyses should be investigated in the future for more causative tracking of health determinants. One example of a recent study that used more sophisticated mathematical underpinnings is Mondal and Shitan⁴⁵, which used path analysis and found a significant association for low and lower-middle income countries between life expectancy and mean years of schooling. In more complex methodological studies there is a tendency for fewer health outcomes, predictors and countries to be analyzed due to data availability problems. Other typical enhancements to the analytical approach are time-series analyses⁴⁶ and multi-level analyses.²⁰

We were struck by the cross-country equity regression results. Although much is known about measuring and monitoring health and coverage inequalities⁴⁷, far less is known about the predictors of the aggregate levels of health inequalities. This is very important for understanding actions to improve health equity and which determinants to monitor. Currently there is little empirical literature using country-level health inequalities metrics^{48,49} as dependent variables. Our paper used gap measures as dependent variables. Predictors were the Gini index, differences in drinking water access, and health systems responsiveness to poorer populations, which were all relevant, but not for all health and service conditions. For global monitoring of SDG-goal 10, covering the reduction of inequalities within countries, it would be useful to know which determinants indicators are most closely linked to health inequalities.

In future, a wider range of indicators could be tested for use in tracking of health determinants as part of the SDGs. We selected what appeared to be feasible indicators, for which for some there

was some available distributional information (i.e. for access to water). But several additional examples of theory-driven indicators were mentioned earlier. Variables already considered in the cross-country literature are female education.⁴⁵ In our analysis we used education overall, but further work would explore female education. Kolves et al.⁵⁰ used the Gini indices, unemployment rates, female participation in the labour force, GDP per capita and divorce rates to predict suicide rates. While Fritzell et al.⁴¹ found child poverty rates and social spending was associated with child mortality. For our dataset of 57 countries, poverty rates were too incomplete to use for all regressions. Another study show that paid maternity leave was also associated with improved immunization coverage.⁴⁸ These studies illustrate the more specific indicators that require further testing, including the importance of policy indicators. Datasets on policy indicators related to the labour market conditions for health may be of specific interest in this regard (see the World Policy Analysis Database: <http://worldpolicycenter.org/>).

A major obstacle to advancing empirical testing of determinants and barriers or facilitators of health services access, like responsiveness, is having both a holistic vision of health, and the available data.³ As part of the SDGs, the United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG) encourages the collection of disaggregated data for monitoring equity across goals – most of which include important health determinants.⁵¹ The follow-up of these recommendations will be very important for any initiatives to track health determinants and their population health and health equity impacts. Investments need to be made to obtain better, disaggregated data about the real sector of the economy, societal well-being and the environment.⁵² Health policy-makers should advocate for better data collection and disaggregation in other sectoral indicators in order to identify common causes across sectors.

CONCLUSIONS

In promoting monitoring of health determinants and related barriers to health service coverage like responsiveness, the health sector will enhance public health promotion, which is necessary for SDG-Goal 3, “attaining healthy life for all at all ages”.⁵ It is only when national health monitoring by the health sector reflects the true intersectoral scope of health that accountability across sectors for actions affecting health will be demanded by the whole-of-society.

REFERENCES

1. Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bull World Health Organ.* 2000;78(6):717-31.
2. World Health Organization. Health systems: improving performance. The world health report 2000. Geneva: WHO; 2000.
3. McKee M. The world health report 2000: 10 years on. *Health Policy Plann.* 2010;25(5):346-8.
4. World Health Organization. Everybody's business: strengthening health systems to improve health outcomes: WHO's framework for action. Geneva: World Health Organization; 2007.
5. World Health Organization, World Bank. Monitoring progress towards universal health coverage: framework, measures and targets. Geneva: World Health Organization; 2014.
6. United Nations Task Team for the Global Thematic Consultation on Health in the Post-2015 Development Agenda. Health in the post-2015 development agenda. Report of the global thematic consultation on health. The world we want. New York: United Nations; 2013.
7. Commission on Social Determinants of Health. Final report. Closing the gap in a generation. Health equity through action on the social determinants of health. Geneva: World Health Organization; 2008.
8. Tanahashi T. Health service coverage and its evaluation. *Bull World Health Organ.* 1978;56(2):295-303.
9. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
10. Andersen R. Revisiting the behavioural model and access to medical care: does it matter. *J Health Soc Behav.* 1995;36(1):1-10.
11. Goddard M. Equity of access to health care services: theory and evidence from the UK. *Social science & medicine.* 2001;53:1149-62.
12. United Nations. Transforming our world: the 2030 agenda for sustainable development In: Assembly UNG, editor. New York: United Nations; 2015.
13. Russell E, Johnson B, Larsen H, Lelinneth M, Novilla B, van Olmen J, et al. Health systems in context: a systematic review of the integration of the social determinants of health within health systems frameworks. *Rev Panam Salud Publica.* 2013;34(6):461-7.
14. Smedby B, Andersen R. International comparisons of health care systems: conceptual and methodological developments over half a century. *Soc Med Tidskr.* 2010;5-6:439-52.
15. Marmor T, Wendt C. Conceptual frameworks for comparing healthcare politics and policy. *Health Policy Plann.* 2012;107:11-20.
16. Muldoon KA, Galway LP, Nakajima M, KanTERS S, Hogg R, Bendavid E, et al. Health system determinants of infant, child and maternal mortality: a cross-sectional study of UN member countries. *Global.* 2011;7(42):e1-10.
17. Hosseinpoor A, Bergen N, Koller T, Prasad A, Schlotheuber A, Valentine NB. Equity-oriented monitoring in the context of universal health coverage. *PLoS Med.* 2014;11(9):e1001727.

Chapter 7. Health System Responsiveness by N Valentine

18. World Health Organization. Measuring and monitoring intersectoral factors influencing equity in universal health coverage (UHC) and health. Summary report of a meeting in Bellagio; 2014 Jul. Report No.: PHE/SDH/2015/1. [online: www.who.int/healthinfo/universal_health_coverage/UHC_Meeting_Bellagio_Mar2014_Report.pdf]
19. Dora C, Haines A, Balbus J, Fletcher E, Adair-Rohani H, Alabaster G. Indicators linking health and sustainability in the post-2015 development agenda. *Lancet*. 2014;385(9965):380-91.
20. Wiysonge CS, Uthman OA, Ndumbe PM, Hussey GD. Individual and Contextual Factors Associated with Low Childhood Immunisation Coverage in Sub-Saharan Africa: a multilevel analysis. *PLOS One*. 2012;7(5):e37905. doi: 10.1371/journal.pone.0037905.
21. Bodkera M, Pisingera C, Toft U, Jørgensena T. The rise and fall of the world's first fat tax. *Health Policy Plann*. 2015;119(6):697-850.
22. Mackenbach JP, Hu Y, Looma CW. Democratization and life expectancy in Europe, 1960-2008. *Soc Sci Med*. 2013;93:166-75.
23. Pickett KE, Wilkinson RG. Income inequality and health: a causal review. *Social science & medicine*. 2015;128:316-26.
24. Valentine NB, Bonsel GJ, Murray CJL. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Qual Life Res*. 2007;16(7):1107-25.
25. Valentine NB, Prasad A, Rice N, Ribone S, Chatterji S. Health systems responsiveness: a measure of the acceptability of health-care processes and systems from the user's perspective. In: Smith P, Mossialos E, Papanicolas I, Leatherman S, editors. *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press; 2010. p. 138-86.
26. Biró A. Copayments, gatekeeping, and the utilization of outpatient public and private care at age 50 and above in Europe. *Health Policy Plann*. 2013;111(1):24-33.
27. Ruger J, Kim H. Out-of-pocket healthcare spending by the poor and chronically ill in the Republic of Korea. *Am J Pub Hlth*. 2007;97(5):804-11.
28. Gakidou E, Cowling K, Lozano R, Murray CJL. Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis. *Lancet*. 2010;376:959-74.
29. Üstün TB, Chatterji S, Mechbal A, Murray CJL, WHS Collaborating Groups. The World Health Surveys. Chapter 58. In: Murray CJL, Evans DB, editors. *Health Systems Performance Assessment: debates, methods and empiricism*. Geneva: World Health Organization; 2003. p.797-808.
30. United Nations Development Agency. *The world development report 2003*. New York: 2003.
31. World Health Organization. *World health statistics 2008*. Geneva; 2008.
32. World Health Organization. *World health statistics 2007*. Geneva; 2007.
33. World Health Organization. *World health statistics 2006*. Geneva; 2006.
34. World Health Organization. *World health statistics 2005*. Geneva; 2005.
35. World Health Survey [Internet]. Geneva:WHO Global Health Observatory. [cited August 2014] Available from: <http://www.who.int/healthinfo/survey/en/>

36. Valentine NB, Verdes-Tennant E, Bonsel GJ. Health systems' responsiveness and reporting behaviour: multilevel analysis of the influence of individual-level factors in 64 countries. *Soc Sci Med.* 2015;138:152-60.
37. Valentine NB, Darby C, Bonsel GJ. Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of "health systems responsiveness" in 41 countries. *Soc Sci Med.* 2008;66(9):1939-50.
38. Gakidou E, Oza S, Vidal F, Li AY, Lee DK, Sousa A. Improving child survival through environmental and nutritional interventions: the importance of targeting interventions toward the poor. *JAMA.* 2007;298:1876-87.
39. van Buuren S. Appendices. Flexible imputation of missing data. Boca Raton, FL.: Chapman & Hall/CRC; 2012. p. 263-7.
40. Rodgers G. Income and inequality as determinants of mortality: an international cross-section analysis. *Pop Stud-J Demog.* 1979;33(2):343-51.
41. Fritzell J, Kangas O, Hertzman J, Blomgren J, Hiilamo H. Cross-temporal and cross-national poverty and mortality rates among developed countries. *J Environ Public Health.* 2013;e915490. doi: 10.1155/2013/915490.
42. Osgood DW. Poisson-based regression analysis of aggregate crime counts. *J Quant Criminol.* 2000;16(1):21-43.
43. DerSarkissian M, Thompson CA, Arah OA. Time series analysis of maternal mortality in Africa. *J Epidemiol Commun H.* 2013;67:992-8.
44. World Health Organization, Ministry of Social Affairs and Health of Finland. The Helsinki statement on health in all policies. World Health Organization and Ministry of Social Affairs, Finland; 2013.
45. Mondal NI, Shitan M. Relative importance of demographic, socioeconomic and health factors on life expectancy in low- and lower-middle-income countries. *J Epidemiol.* 2014;24(2):117-24.
46. Chuang Y, Sung P, Chao H, Bai C, Chang C. A longitudinal ecological study of the influences of political, economic, and health services characteristics on under-five mortality in less-developed countries *Health Place.* 2013;23 (September):111-21.
47. Hosseinpoor AR, Bergen N, Schlotheuber A. Promoting health equity: WHO health inequality monitoring at global and national levels. *Glob Health Action.* 2015;8e29034.doi: 10.3402/gha.v8.29034.
48. Hajizadeha M, Heymann J, Strumpf E, Harper S, Nadi A. Paid maternity leave and childhood vaccination uptake: longitudinal evidence from 20 low-and-middle-income countries. *Soc Sci Med.* 2015 (September);140:104-17.
49. Backhans M, Burström B, Ponce de Leon A, Marklund S. Is gender policy related to the gender gap in external cause and circulatory disease mortality? A mixed effects model of 22 OECD countries 1973-2008. *BMC Public Health.* 2012;12:e969. doi: 10.1186/1471-2458-12-969.
50. Kolves K, Milner A, Varn P. Suicide rates and socioeconomic factors in Eastern European countries after the collapse of the Soviet Union: trends between 1990 and 2008. *Social Health Ill.* 2013;35(6):956-70.

Chapter 7. Health System Responsiveness by N Valentine

51. Independent Expert Advisory Group (IEAG) on a Data Revolution for Sustainable Development. A world that counts. Mobilising the data revolution for sustainable development. New York:United Nations; 2014 (November).
52. Commission on the Measurement of Economic Performance and Social Progress. Report of the Commission on the Measurement of Economic Performance and Social Progress. Paris; 2008.

Appendix 7.1. Countries in regression analyses: outcomes variables for health status and coverage

| Country | Regression 1: Maternal mortality per 100,000 lbs - 2005 | Regression 2: Child mortality per 1000 live births - 2005 | Regression 3: TB cause of death per 100,000 - 2004 | Regression 4: Percentage of births attended by skilled health personnel - 2000-06 (2008 WHS) | Regression 5: Percentage of population with coverage with one dose of measles vaccination in the first year of life - 2003 (WHS 2005) | Regression 6: Percentage of women receiving a pap smear(2000-06) (WHS 2008, 58 countries) |
|------------------------|--|--|---|---|--|--|
| Bangladesh | 570 | 69 | 4 | 20 | 77 | 0 |
| Bosnia and Herzegovina | 3 | 17 | 6 | 100 | 84 | 40 |
| Brazil | 110 | 35 | 7 | 97 | 99 | 72 |
| Burkina Faso | 700 | 207 | 54 | 54 | 76 | 5 |
| Chad | 1500 | 200 | 82 | 14 | 61 | 6 |
| China | 45 | 37 | 16 | 98 | 84 | 21 |
| Comoros | 400 | 73 | 7 | 62 | 63 | 8 |
| Congo | 740 | 108 | 70 | 83 | 50 | 23 |
| Cote d'Ivoire | 810 | 193 | 104 | 57 | 56 | 7 |
| Croatia | 7 | 7 | 6 | 100 | 95 | 65 |
| Czech Republic | 4 | 5 | 1 | 100 | 99 | 73 |
| Democratic Republic | 660 | 91 | 25 | 19 | 42 | 3 |
| Dominican Republic | 150 | 35 | 16 | 96 | 79 | 66 |
| Ecuador | 210 | 27 | 25 | 80 | 99 | 45 |
| Estonia | 25 | 8 | 6 | 100 | 95 | 53 |
| Ethiopia | 720 | 169 | 79 | 6 | 52 | 1 |
| Finland | 7 | 4 | 1 | 100 | 97 | 67 |
| France | 8 | 5 | 1 | not available | 86 | 75 |
| Georgia | 66 | 45 | 13 | 92 | 73 | 13 |
| Ghana | 560 | 95 | 50 | 50 | 80 | 3 |
| Hungary | 6 | 9 | 3 | 100 | 99 | 65 |
| India | 450 | 87 | 30 | 47 | 67 | 3 |
| Ireland | 1 | 6 | 1 | 100 | 78 | 39 |
| Israel | 4 | 6 | 1 | not available | 95 | 45 |
| Kazakhstan | 140 | 73 | 20 | 100 | 99 | 79 |
| Kenya | 560 | 123 | 133 | 42 | 72 | 4 |

Appendix 7.1 continued

| Country | Regression 1: Maternal mortality per 100,000 lbs - 2005 | Regression 2: Child mortality per 1000 live births - 2005 | Regression 3: TB cause of death per 100,000 - 2004 | Regression 4: Percentage of births attended by skilled health personnel - 2000-06 (2008 WHS) | Regression 5: Percentage of population with coverage with one dose of measles vaccination in the first year of life - 2003 (WHS 2005) | Regression 6: Percentage of women receiving a pap smear(2000-06) (WHS 2008, 58 countries) |
|----------------------|--|--|---|---|--|--|
| Latvia | 10 | 13 | 10 | 100 | 99 | 3 |
| Luxembourg | 12 | 4 | 1 | 100 | 91 | 82 |
| Malawi | 1100 | 178 | 97 | 54 | 77 | 3 |
| Malaysia | 62 | 7 | 16 | 100 | 92 | 30 |
| Mali | 970 | 220 | 73 | 41 | 68 | 5 |
| Mauritania | 820 | 184 | 60 | 53 | 71 | 4 |
| Mauritius | 15 | 17 | 11 | 99 | 94 | 13 |
| Mexico | 60 | 28 | 4 | 94 | 96 | 64 |
| Myanmar | 380 | 106 | 20 | 57 | 75 | 1 |
| Namibia | 210 | 65 | 81 | 76 | 70 | 13 |
| Nepal | 830 | 82 | 23 | 19 | 75 | 3 |
| Norway | 7 | 4 | 1 | 84 | 73 | 73 |
| Pakistan | 320 | 103 | 40 | 61 | 3 | 3 |
| Paraguay | 150 | 29 | 12 | 91 | 53 | 53 |
| Philippines | 230 | 36 | 48 | 80 | 10 | 10 |
| Portugal | 11 | 6 | 4 | 96 | 59 | 59 |
| Russian Federation | 28 | 16 | 21 | 96 | 78 | 78 |
| Senegal | 980 | 137 | 52 | 96 | 11 | 11 |
| Slovakia | 6 | 8 | 3 | 99 | 59 | 59 |
| South Africa | 400 | 66 | 134 | 83 | 6 | 6 |
| Spain | 4 | 5 | 2 | 97 | 60 | 60 |
| Sri Lanka | 58 | 15 | 9 | 99 | 2 | 2 |
| Swaziland | 390 | 153 | 269 | 94 | 62 | 62 |
| Sweden | 3 | 4 | 0.5 | 94 | 70 | 70 |
| Tunisia | 100 | 24 | 2 | 90 | 10 | 10 |
| Ukraine | 18 | 20 | 16 | 99 | 34 | 34 |
| United Arab Emirates | 37 | 8 | 2 | 94 | 12 | 12 |
| Uruguay | 20 | 15 | 3 | 95 | 62 | 62 |

Appendix 7.1 continued

| Country | Regression 1: Maternal mortality per 100,000 lbs - 2005 | Regression 2: Child mortality per 1000 live births - 2005 | Regression 3: TB cause of death per 100,000 - 2004 | Regression 4: Percentage of births attended by skilled health personnel - 2000-06 (2008 WHS) | Regression 5: Percentage of population with coverage with one dose of measles vaccination in the first year of life - 2003 (WHS 2005) | Regression 6: Percentage of women receiving a pap smear(2000-06) (WHS 2008, 58 countries) |
|----------|--|--|---|---|--|--|
| Viet Nam | 150 | 23 | 22 | 93 | 7 | 7 |
| Zambia | 830 | 182 | 138 | 84 | 3 | 3 |
| Zimbabwe | 880 | 126 | 131 | 80 | 9 | 9 |
| Average | 308 | 63 | 36 | 84 | 31 | 31 |
| Std Dev | 368 | 66 | 50 | 15 | 29 | 29 |
| Min | 1 | 4 | 1 | 42 | 0 | 0 |
| Max | 1500 | 220 | 269 | 99 | 82 | 82 |

Appendix 7.2 Comparing univariate distributions for skewness and Kurtosis statistics: example for health outcomes

| Variables | Distribution | | |
|--|----------------------|----------|----------|
| | Variable form logged | Skewness | Kurtosis |
| Maternal mortality rate per 100,000 live births (2005) | No | 1.15 | 3.48 |
| Natural log of maternal mortality rate | Yes | -0.35 | 1.78 |
| Maternal mortality counts | No | 4.79 | 25.24 |
| Natural log of maternal mortality counts | Yes | -0.17 | 1.81 |
| Under five child mortality per 1000 live births (2005) | No | 0.98 | 2.65 |
| Natural log of child mortality rate | Yes | -0.12 | 1.64 |
| Child mortality counts | No | 4.77 | 27.11 |
| Natural log of child mortality counts | Yes | -0.04 | 2.07 |
| TB cause of death per 100,000 (2004) | No | 2.32 | 9.66 |
| Natural log of TB cause | Yes | -0.19 | 1.99 |

PART III

Using responsiveness measures in the Netherlands' sub-system of perinatal care

CHAPTER 8

Validity of a questionnaire measuring the World Health Organization concept of health system responsiveness with respect to perinatal services in the Dutch obstetric care system



J. van der Kooy, N.B. Valentine, E. Birnie, M.Vujkovic, J.P. de Graaf,
S. Denktas, E.A.P Steegers, G.J. Bonsel

BMC Health Services Research. 2014; 14:e622

ABSTRACT

Background. The concept of responsiveness, introduced by the World Health Organization (WHO), addresses non-clinical aspects of health service quality that are relevant regardless of provider, country, health system or health condition. Responsiveness refers to “aspects related to the way individuals are treated and the environment in which they are treated” during health system interactions. This paper assesses the psychometric properties of a newly developed responsiveness questionnaire dedicated to evaluating maternal experiences of perinatal care services, called the Responsiveness in Perinatal and Obstetric Health Care Questionnaire (ReproQ), using the 8-domain WHO concept.

Methods. The ReproQ was developed between October 2009 and February 2010 by adapting the WHO responsiveness questionnaire items to the perinatal care context. The psychometric properties of feasibility, construct validity, and discriminative validity were empirically assessed in a sample of Dutch women 2 weeks post partum.

Results. A total of 171 women consented to participation. Feasibility: the interviews lasted between 20 and 40 minutes and the overall missing rate was 8%. Construct validity: mean Cronbach's alphas for the antenatal, birth and postpartum phase were: 0.73 (range 0.57-0.82), 0.84 (range 0.66-0.92), and 0.87 (range 0.62-0.95) respectively. The item-own scale correlations within all phases were considerably higher than most of the item-other scale correlations. Within the antenatal care, birth care and post partum phases, the 8 factors explained 69%, 69%, and 76% of variance respectively. Discriminative validity: overall responsiveness mean sum scores were higher for women whose children were not admitted. This confirmed the hypothesis that dissatisfaction with health outcomes is transferred to their judgement on responsiveness of the perinatal services.

Conclusions. The ReproQ interview-based questionnaire demonstrated satisfactory psychometric properties to describe the quality of perinatal care in the Netherlands, with the potential to discriminate between different levels of quality of care. In view of the relatively small sample, further testing and research is recommended.

INTRODUCTION

The debate on the organization of perinatal care in the Netherlands has intensified over recent years. The Dutch perinatal health care system has come under pressure since the national perinatal mortality rates were shown to be among the highest in Europe.¹ This system can be regarded as a sequential chain of health care services, each dedicated to a different phase of the perinatal experience: antenatal care, birth and post partum care. Antenatal, birth-related and post partum care are provided by different caregivers with different responsibilities, for different risk groups, and in different settings. In the Netherlands independently operating community midwives provide care for low-risk pregnant women (primary health care) while gynaecologists provide in-hospital care for high-risk women (secondary care). All women receive post partum care by a community midwife.

The performance of perinatal care is often judged by its endpoints such as clinical outcomes and costs. However, quality of care literature supports the view that non-clinical aspects of quality are important too, and affect clinical outcomes.²⁻⁴ Better service quality is thought to increase compliance with medical treatment and to improve the transfer of information and appropriate utilization of health care.⁵⁻⁸ Governments of Western countries increasingly acknowledge the importance of the non-clinical aspects of quality of care and incorporate these when the provision of care is monitored.^{9,10}

Sofar, no attempts have been made to evaluate the non-clinical aspects of the Dutch obstetric care system such that not only the heterogeneity in the quality with respect to different perinatal services is identified, but also that international comparisons with other obstetric care systems are possible.¹¹⁻¹³ The concept of responsiveness, introduced by the World Health Organization (WHO) in 2000, seems apt to this task as it was specifically developed to refer to patients' experiences when interacting with health care providers. The concept reviews 8 predefined domains relevant to non-clinical aspects of service quality regardless of provider, country, health system or health condition. Responsiveness is defined as "aspects related to the way individuals are treated and the environment in which they are treated during health system interactions" (page 574), encompassing the notions of both non-clinical quality and patient experience.¹⁴ The concept of responsiveness excludes the financial and clinical domains and focuses on a set of non-clinical domains that reflect respect for human dignity and the client orientation of the care process and setting. While these domains may influence health outcomes, health outcomes are not part of the responsiveness concept. The relevance of an independent set of non-clinical domains to health systems performance is supported by the discipline of medical ethics and in human rights law, which argues that responsiveness features of a health system are important in their own right.¹⁴⁻¹⁶ The concept of responsiveness aims to support measurement of ser-

vice quality in an internationally comparable way and to enable quantitative trade-offs between non-clinical aspects of service quality and clinical outcomes.¹⁴ The concept of responsiveness aims to capture information on the non-clinical quality of the patient's actual experience in contrast to patient satisfaction questionnaires. Literature has shown that expectations may strongly influence patient satisfaction, which makes international comparisons of non-clinical service quality challenging since expectations are in turn influenced by economic and political influences.¹⁷⁻²⁰

Adopting the responsiveness concept, the Responsiveness in Perinatal and Obstetric Health Care Questionnaire (ReproQ), was developed by adapting the existing generic World Health Survey questionnaire responsiveness module into a questionnaire dedicated to maternal experiences during perinatal care. The aim of this study is to investigate the psychometric properties of the ReproQ.

METHODS

Questionnaire

The WHO developed a survey, which was administrated between 2000-2001 under the auspices of the Multi-Country Survey Study on Health and Health Systems Responsiveness (MCS Study) and again in 2002-03 under the World Health Survey (WHS).^{14,21} The concept of responsiveness, containing 8 domains, was identified in WHO's review of the patient satisfaction and quality of care literature.¹⁵

Several questionnaires and related studies relevant to responsiveness domains were used, such as the Community Tracking Study²², Picker Survey²³, QUOTE study²⁴ and the CAHPS (Consumer Assessment of Health Plans Study).²⁵

Although there are many overlapping aspects with patient satisfaction questionnaires, the concept of responsiveness is different on several points; where patient satisfaction generally covers both medical and non-medical aspects of care, responsiveness focuses only on the non-clinical enhancing aspects of the health system. Where patient satisfaction represents a complex mixture of perceived need, individually determined expectations and experience of care, responsiveness evaluates individual's perceptions of the health system against 'legitimate' expectations – referring to standards that can be applied everywhere, or 'universally'.¹⁵

The ReproQ was developed between October 2009 and February 2010, and its questions were derived from these WHO questionnaires.

The ReproQ questionnaire was developed to assess the responsiveness outcomes of perinatal health care system in the Netherlands and is based on the same 8 domains identified in WHO's review, i.e. dignity, autonomy, confidentiality, communication, prompt attention, social consideration (labelled initially as *Access to Social Support* or *Access to Family and Community Support*), quality of basic amenities, and choice (and continuity).

These domains are claimed to be of importance in all health systems, during any client-system interaction (including personal and non-personal health services) and for the population's interaction with insurers and other administrative arms of the health system. While it is recognized that persons may differ regarding the relative importance of each domain, and that specific domains may be of extra relevance in particular health care interactions, it is assumed that the quality of any interaction is sufficiently covered by these 8 domains.¹⁴

The ReproQ asks the same questions for the three phases of perinatal care: antenatal phase (the period from the onset of pregnancy until the onset of delivery), birth phase (actual delivery) and post partum phase (covering the first 10 days after childbirth). Rather than pointing to a single event, or the last visit (as in the WHS), the questionnaire we selected to focus questions on women's judgments for all antenatal visits as one done for the MCS Study. The 'last visit' approach has better recall but is easily biased by a particular incident. We wanted to review the experience as a whole and thought the multiple visit approach more suited to this. A similar argument applied to the decision to focus postnatal maternity care questions on all visits. For the birth phase, it seemed appropriate to focus questions on the single event of 'delivery'. Within this framework, the setting and professional items were adapted to the perinatal care in the obstetric care system (e.g., "doctor" was translated into "midwife" or "gynaecologist"). If two different health care professionals could be involved (e.g., "midwife" and "nurse" during delivery), similar questions within each domain were repeated for each health care professional separately.

Each phase was covered by the above mentioned 8 domains, with 2-7 items per domain. The standardized response mode consisted of 5 options: "very good", "good", "moderate", "bad", and "very bad". The ReproQ consisted of 104 questions on responsiveness (25 antenatal, 40 birth, 39 postpartum phase) and 29 questions for maternal and health care characteristics.

Questions from the WHO questionnaire were translated into Dutch according to a predefined protocol. First, questionnaires were translated by the research team. Expert meetings consisting of gynaecologists, midwives, nurses, public health experts and researchers were held to judge the translation and comprehensiveness of the item list. Many among these professionals had working experience in English speaking countries. Next, backward translation of each question

was then performed and comparison was made with the original English questionnaire. Improvements were made and final consensus was reached on each question.

The completeness of domains was judged in terms of being comprehensive (are all non-clinical areas covered, which clients and professionals put forward either as positive experiences or negatively as complaint), and in terms of being balanced (have all domains included given equal importance). For each domain the candidate pool of items was checked whether each item fitted to the domain definition sufficiently. As this could differ per phase, this was discussed for each phase separately (e.g., the item “quality of the food” during antenatal visits was excluded). Finally we asked the experts to check whether all the domains would remain valid under ongoing and anticipated organizational changes in perinatal care. All stakeholders agreed that in the final list the stated requirements were met.

Finally, 6 primiparous and multiparous pregnant women were invited to judge the feasibility of the draft version of the questionnaire. Since we adopted an existing concept and adapted questions from an extensively studied source questionnaire towards a perinatal context, we invited the judgment of these 6 women in the final stage. They were first asked to conduct a brainstorm on important non-clinical aspects of perinatal care. Next, the ReproQ was evaluated to see whether its domains covered these issues. All items were discussed separately including their meaning and understandability. The original domain structure proved to be comprehensive, as judged by the stakeholders. Small textual improvements were made in the item questions as a result of this meeting (Table 8.1).

Table 8.1 The 8 domains with the items given for the antenatal phase

| Domain | Question / Item |
|---|---|
| Dignity | Were physical examinations and treatments done in a way that respected your privacy? Did the examination rooms ensure your privacy? Were you treated with respect by your health care provider? |
| Autonomy | How well were you involved in making decisions regarding your examinations or treatments? Were you able to refuse examinations or treatments? Were you asked permission before testing or starting treatment? |
| Confidentiality of information | Were consultations carried out in a manner that protected your confidentiality? Was confidentiality kept on the information provided by you? Was your medical record kept confidential? |
| Communication | How well were things explained by your health care provider in a way you could understand? Was written information provided in such a way you could understand? Were you encouraged to ask questions about your health problems, treatment and care? Were you given time to ask questions about your health problem or treatment? Was information on the health service's contact, location and parking information clear to you? |
| Prompt attention | How well did you receive prompt attention at your health service? How did you experience the waiting time after you asked for help? How well was the accessibility by phone? How do you rate the travel time to your health service? |
| Social consideration | Did the health care provider facilitate the support of your relatives and friends? Was the home situation taken into consideration when planning an appointment? |
| Quality of basic amenities | How do you rate the quality of the hygiene of the toilets? How do you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanliness? |
| Choice and continuity (of health care provider) | Were you able to choose your own health care provider? Were you able to use other health care services other than the one you usually went to? How well was the continuity of care by one health care provider? Were you able to choose your own place of delivery? |

Study population and data collection

Study approval was granted by the Medical Ethical Committee, Erasmus Medical Centre, Rotterdam, the Netherlands, no MEC2012207. To investigate the psychometric properties of responsiveness questions for each phase, women were recruited from 3 midwifery practices in Rotterdam, the Netherlands, between February 2010 and March 2011 (all women, regardless of their health utility, received post partum care by a community midwife in the Netherlands). Women or their partners were required to speak and understand Dutch sufficiently. Written informed consent was obtained.

The survey was administered in the form of face-to-face interviews 2 weeks after delivery. Face-to-face interviews were chosen since this method enhances participation, in particular by those with low education and migrants, and since this method was also chosen in the WHO survey. A randomly selected subset of women was invited by their own midwife for study participation. The interview took 30 minutes face-to-face with an independent interviewer. The interviews were conducted by 10 trained independent interviewers and usually performed at the respondent's home. Each interview covered all three phases of the maternal perinatal experience. Interviewees were invited to respond to all questions, yet never forced to. Of the different interview modes, face-to-face interviews were chosen as this mode reduces non-response bias. The face-to-face mode was also the preferred one used for a large number of the MCS Study countries and in the World Health Survey.

Data handling and analysis

Records were regarded 'missing at the record level' if all scores of all phases were missing. If women had responded partially, the responses were evaluated per phase. If all the items of one phase were missing, this record was excluded from the analysis of that phase. This implies that occasionally respondents were excluded from one phase while they were included in the analysis of other phases. Missing items were excluded from analysis.

We investigated the responsiveness questions' psychometric properties stratified for the antenatal phase (the period from the onset of pregnancy until the onset of delivery), the birth phase (actual delivery) and post partum phase (covering the first 10 days after childbirth). The data were analysed with Statistical Package of Social Sciences version 20.0 for Windows (IBM Corp. Released 2011).

Sumscores

Unweighted sumscores per domain were calculated and transformed into 1-10 scale scores to enhance comparability among domains with different numbers of items. Transformation was done as $\text{score} = 1 + 9 * \frac{([\text{sumscore} - \text{lowest sum possible}])}{([\text{largest sum possible} - \text{lowest sum possible}]})$; e.g., a domain that contains 3 items each with a 5-point response mode, displays a possible score range from 3 to 15. The transformed sumscore would then be $1 + 9 * \frac{([\text{sumscore} - 3])}{[15 - 3]}$. If sumscore in an individual were 11, her transformed score would be $1 + 9 * \frac{[11 - 3]}{[15 - 3]} = 1 + 9 * \frac{8}{12} = 7$. This transformation procedure was repeated for each domain in each phase separately.

Psychometric tests

The following psychometric properties of the ReproQ were evaluated: feasibility, construct validity, and discriminative validity. Feasibility was expressed as rates of missing items per domain. The literature provides little indication of acceptable survey response rates or inappropriate non-response rates. We selected missing item rates below 20% as acceptable as done for another study.²¹ In addition, we compared missing rates per item for each phase to identify problematic single items. Furthermore, we compared missing rates per domain by age, education, race, communication and health utilization to check for biases by social groups.

Scores per domain, expressed as transformed 1-10 scale scores, and scores per item, given in 1-5 scale scores, were described in terms of mean, SD, range, floor and ceiling effects, and percentiles.

Reliability was assessed as internal consistency by using Cronbach's alpha. Amidst varying standards in the literature, we considered 0.70 to be an acceptable alpha coefficient.²⁶ Average inter-item, average item-own scale and average item-other scale correlation were assessed with standardized correlation coefficients, with acceptable correlations defined as Pearson's correlation coefficients (r) > 0.40.²⁷ We expected higher average inter-item and average inter-own scale correlations compared to average inter-other scale correlations.

Discriminative validity was assessed by comparing subgroups expected to differ in terms of responsiveness. It was hypothesized that women whose child was not admitted to the hospital would report better responsiveness outcomes than women whose child was hospitalized. The rationale behind this hypothesis is that women with less good clinical outcomes would be more critical on the non-clinical aspects of care given. Differences in overall mean sum scores (adding all domains) were calculated and tested with Student t-tests per phase.

Construct validity was assessed as the domain structure of factor loadings obtained with exploratory factor analysis using maximum likelihood method with oblique promax rotation of factor loadings, extracting 8 (fixed) factors. This was done to explore whether the original domain structure relevant to the generic responsiveness concept was present after adapting the responsiveness concept to perinatal services.

RESULTS

Of a total of 274 women who were identified for study participation, 94 women could not be reached or they declined the invitation; many women could not be reached using the cell phone number they had provided; we were unable to differentiate with limited means whether they refused the call, changed phone number, or had provided the wrong number. Other reasons for non-participation included lack of time, and feeling unease at having a stranger visit their home. One hundred and eighty women (66%) agreed to be interviewed. Of these, 7 interviews (7/180, 4%) were cancelled and 2 interviews (2/180, 1%) were discontinued because the respondents did not speak Dutch with sufficient fluency and no translator could be made available. The remaining 171 interviews were used for analysis. The interviews took between 20 and 40 minutes. Table 8.2 describes the characteristics of the participants.

The mean maternal age was 31 years (95% CI 30.3–31.7). The majority of mothers were primiparous (57%). A substantial proportion of mothers was of non-Dutch origin (43%), or lived in underprivileged neighbourhoods (51%). Few had low education (4%) or were single (18%). Approximately 11% spoke weak/poor Dutch as judged by the interviewer. Referral to gynaecologists had occurred in approximately 55% of women. Post partum hospital admission had occurred in 26 (15%) of all newborns.

Table 8.3 describes the missing rates per domain, for each phase separately. The table also describes the maximum missing rate per item for that domain. The results for 4 women, with no response in the birth phase, were excluded.

The average item missing rate over all phases was 8% (1,349 out of 17,624 questions). Missing rates per domain were all below the predefined threshold of 20%. Average missing rates across domains were highest in the birth phase (8%). Maximum item missing rates per domain ranged from 1.8% to 11.1% for the antenatal phase, from 5.3% to 31.6% for the birth phase, and from 5.3% to 14.6% for the post partum phase (see Appendix 8.1 for detailed description of all items). The highest item missing rate was for 2 questions relevant to the birth phase: “able to be referred to a medical specialist during birth care” (31.6%) and, “consideration of home situation when planning appointments / examinations during birth care” (22.8%). Item missing pertained

mainly to the birth care phase and rates were higher among women of Dutch origin. There were no differences in missing rates by age, educational level and health utilization.

Table 8.2 Characteristics of the participants

| Variable | Number | Percent |
|---|--------|---------|
| Maternal Age^a | | |
| <19 years | 3 | 2% |
| 20-25 years | 15 | 9% |
| 25-34 years | 119 | 70% |
| >35 years | 33 | 19% |
| Missing | 1 | 1% |
| Parity | | |
| Primiparous | 97 | 57% |
| Multiparous | 74 | 43% |
| Ethnic background | | |
| Dutch | 94 | 55% |
| Non Dutch | 74 | 43% |
| Education | | |
| Low | 6 | 4% |
| Middle | 75 | 44% |
| High | 90 | 53% |
| Marital status | | |
| Single | 30 | 18% |
| relationship/married | 141 | 82% |
| Neighbourhood | | |
| privileged neighbourhood | 84 | 49% |
| underprivileged neighbourhood | 87 | 51% |
| Proficiency (speaking) Dutch | | |
| good/excellent | 152 | 89% |
| weak/poor | 18 | 11% |
| Missing | 1 | 1% |
| Care process | | |
| start antenatal care with midwife, not referred | 61 | 36% |
| start antenatal care with midwife, referred during antenatal phase to gynaecologist | 37 | 22% |
| start antenatal care with midwife, referred during birth phase | 57 | 33% |
| start antenatal care with gynaecologist | 16 | 9% |
| Hospital admission of child | | |
| No admission | 145 | 85% |
| Admission | 26 | 15% |

^a mean age 31 (95%CI 30.0-31.7)

Table 8.3 Missing item values and the maximum percentage missing per item, for each domain and perinatal phase

| Domain | Antenatal Phase (n=171) ^a | | | | Birth Phase (n=167) ^a | | | | Post Partum Phase (n=171) ^a | | | | Total (n=509) | | |
|----------------------------|--------------------------------------|-------------------|-------------------------------|------------------------|----------------------------------|-------------------|-------------------------------|------------------------|--|-------------------|-------------------------------|------------------------|---------------|------------------------|-----|
| | Total items | Number per domain | Missing per item, maximum (%) | Missing per domain (%) | Total items | Number per domain | Missing per item, maximum (%) | Missing per domain (%) | Total items | Number per domain | Missing per item, maximum (%) | Missing per domain (%) | Number | Missing per domain (%) | |
| Dignity | 513 | 7 | 1% | 1.8% | 835 | 21 | 3% | 5.3% | 855 | 55 | 6% | 10.5% | 2203 | 83 | 4% |
| Autonomy | 513 | 57 | 11% | 11.1% | 501 | 57 | 11% | 14.0% | 855 | 112 | 13% | 15.8% | 1869 | 226 | 12% |
| Confidentiality | 513 | 28 | 5% | 7.0% | 1002 | 88 | 9% | 11.1% | 1026 | 91 | 9% | 11.7% | 2541 | 207 | 8% |
| Communication | 855 | 35 | 4% | 6.4% | 1002 | 49 | 5% | 11.7% | 1026 | 65 | 6% | 14.6% | 2883 | 149 | 5% |
| Prompt attention | 684 | 28 | 4% | 8.8% | 1169 | 111 | 9% | 18.1% | 684 | 58 | 8% | 15.2% | 2537 | 197 | 8% |
| Social consideration | 342 | 16 | 5% | 4.1% | 501 | 65 | 13% | 22.8% | 855 | 60 | 7% | 12.3% | 1698 | 141 | 8% |
| Quality of basic amenities | 342 | 8 | 2% | 1.8% | 501 | 45 | 9% | 11.1% | 513 | 35 | 7% | 5.3% | 1356 | 88 | 6% |
| Choice and continuity | 513 | 45 | 9% | 10.5% | 1169 | 130 | 11% | 31.6% | 855 | 83 | 10% | 14.6% | 2537 | 258 | 10% |
| Total | 4275 | 224 | 5% | | 6680 | 566 | 8% | | 6669 | 559 | 8% | | 17624 | 1349 | 8% |

^a For each phase separately records of non-responders were excluded

Table 8.4 Mean (SD) transformed score, range, percentage floor and ceiling response, and Cronbach's α for each domain and phase

| Domain | No. of items | Mean | SD | Range | % Floor | % Ceiling | 25th %tile | 50th %tile | 75th %tile | Cronbach's α | |
|-----------------------------------|--------------|------|-----|-------|---------|-----------|------------|------------|------------|---------------------|------|
| Dignity | | | | | | | | | | | |
| Antepartum Phase | 3 | 8.4 | 1.1 | 5.5 | 10.0 | 0.0% | 21.6% | 7.8 | 7.8 | 9.3 | 0.73 |
| Birth Phase | 5 | 8.1 | 1.1 | 1.1 | 10.0 | 0.0% | 11.7% | 7.8 | 7.8 | 9.1 | 0.86 |
| Post Partum Phase | 5 | 7.9 | 1.3 | 3.3 | 10.0 | 0.0% | 12.3% | 7.8 | 7.8 | 8.2 | 0.87 |
| Autonomy | | | | | | | | | | | |
| Antepartum Phase | 3 | 7.8 | 1.2 | 3.3 | 10.0 | 0.0% | 8.2% | 7.0 | 7.8 | 8.5 | 0.73 |
| Birth Phase | 3 | 7.7 | 1.4 | 1.4 | 10.0 | 0.0% | 8.8% | 7.7 | 7.8 | 7.8 | 0.87 |
| Post Partum Phase | 5 | 7.5 | 1.7 | 1.9 | 10.0 | 0.0% | 0.6% | 7.3 | 7.8 | 7.8 | 0.94 |
| Confidentiality | | | | | | | | | | | |
| Antepartum Phase | 3 | 8.0 | 1.1 | 4.0 | 10.0 | 0.0% | 14.0% | 7.8 | 7.8 | 8.5 | 0.82 |
| Birth Phase | 6 | 7.8 | 1.4 | 1.4 | 10.0 | 0.0% | 12.3% | 7.8 | 7.8 | 7.8 | 0.78 |
| Post Partum Phase | 6 | 7.7 | 1.4 | 1.8 | 10.0 | 0.0% | 13.5% | 7.4 | 7.8 | 7.8 | 0.94 |
| Communication | | | | | | | | | | | |
| Antepartum Phase | 5 | 7.7 | 1.2 | 3.3 | 10.0 | 0.0% | 5.3% | 7.3 | 7.8 | 8.2 | 0.80 |
| Birth Phase | 6 | 7.8 | 1.3 | 1.3 | 10.0 | 0.0% | 9.9% | 7.4 | 7.8 | 8.1 | 0.92 |
| Post Partum Phase | 6 | 7.6 | 1.7 | 1.0 | 10.0 | 0.6% | 11.7% | 7.4 | 7.8 | 8.1 | 0.95 |
| Prompt attention | | | | | | | | | | | |
| Antepartum Phase | 4 | 7.1 | 1.4 | 1.0 | 10.0 | 0.6% | 2.3% | 6.6 | 7.2 | 7.8 | 0.67 |
| Birth Phase | 7 | 7.7 | 1.3 | 1.3 | 10.0 | 0.0% | 7.0% | 7.1 | 7.8 | 8.4 | 0.83 |
| Post Partum Phase | 4 | 7.7 | 1.7 | 1.0 | 10.0 | 0.6% | 12.9% | 7.2 | 7.8 | 8.9 | 0.89 |
| Social consideration | | | | | | | | | | | |
| Antepartum Phase | 2 | 7.1 | 1.8 | 1.0 | 10.0 | 0.6% | 8.2% | 5.5 | 7.8 | 7.8 | 0.76 |
| Birth Phase | 3 | 7.6 | 1.6 | 1.6 | 10.0 | 0.6% | 11.1% | 7.0 | 7.8 | 7.8 | 0.87 |
| Post Partum Phase | 5 | 7.8 | 1.4 | 3.3 | 10.0 | 0.0% | 8.2% | 7.3 | 7.8 | 8.7 | 0.84 |
| Quality of basic amenities | | | | | | | | | | | |
| Antepartum Phase | 2 | 7.5 | 1.4 | 3.3 | 10.0 | 0.0% | 10.5% | 6.6 | 7.8 | 7.8 | 0.57 |
| Birth Phase | 3 | 7.6 | 1.4 | 1.4 | 10.0 | 0.0% | 8.2% | 7.0 | 7.8 | 8.5 | 0.66 |
| Post Partum Phase | 3 | 7.4 | 1.5 | 1.8 | 10.0 | 0.0% | 6.4% | 7.0 | 7.8 | 7.8 | 0.62 |
| Choice and continuity | | | | | | | | | | | |
| Antepartum Phase | 3 | 7.3 | 1.7 | 1.0 | 10.0 | 0.6% | 7.0% | 6.3 | 7.8 | 7.8 | 0.77 |
| Birth Phase | 7 | 7.2 | 1.5 | 1.5 | 10.0 | 0.0% | 5.3% | 6.5 | 7.6 | 7.8 | 0.88 |
| Post Partum Phase | 5 | 7.1 | 1.7 | 1.0 | 10.0 | 0.6% | 7.0% | 6.4 | 7.8 | 7.8 | 0.89 |

Table 8.5 Promax rotated factor solution for the birth phase

| Factor Name | Confidentiality | Choice and continuity | Dignity | Prompt attention | Autonomy | Communication | Quality of basic amenities | Social consideration | Unique variance |
|--|-----------------|-----------------------|-------------|------------------|--------------|---------------|----------------------------|----------------------|-----------------|
| Respect shown during examinations (midwife) | .032 | .096 | .670 | -.004 | -.053 | -.004 | -.178 | .243 | .305 |
| Examination room suitable to provide privacy | .090 | .018 | .792 | .018 | -.163 | -.221 | .137 | -.013 | .277 |
| Treated with respect (midwife) | -.103 | -.138 | .871 | .037 | -.059 | .001 | .057 | .071 | .310 |
| Respect shown during examinations (nurse) | -.004 | .031 | .768 | -.051 | .092 | .193 | -.095 | -.061 | .267 |
| Treated with respect (nurse) | -.056 | -.103 | .568 | .032 | .017 | .244 | .067 | .029 | .265 |
| Involved in making a decision regarding your examinations or treatments | -.163 | .041 | .065 | -.058 | .895 | -.051 | .010 | .080 | .279 |
| Able to refuse examinations or treatments | .011 | -.029 | -.162 | .044 | 1.009 | .059 | -.102 | -.006 | .242 |
| Asked permission before testing or starting treatment | .059 | .187 | -.033 | -.046 | .693 | -.067 | -.046 | .003 | .377 |
| Protecting your confidentiality during consultations (midwife) | .635 | .055 | .135 | -.003 | .062 | -.048 | .100 | .039 | .234 |
| Confidentiality kept on provided information (midwife) | .773 | .076 | .099 | -.098 | .055 | .014 | .078 | -.064 | .179 |
| Confidentiality of patients' medical records preserved (midwife) | .843 | .063 | .086 | -.117 | .065 | -.020 | -.070 | .033 | .161 |
| Protecting your confidentiality during consultations (nurse) | .548 | -.081 | -.051 | .091 | -.192 | -.126 | .071 | .156 | .584 |
| Confidentiality kept on provided information (nurse) | .955 | -.079 | -.066 | .144 | -.108 | .113 | -.023 | -.050 | .156 |
| Confidentiality of patients' medical records preserved (nurse) | .872 | .012 | -.160 | .099 | .046 | .089 | .003 | .016 | .140 |
| Information clearly explained (midwife) | .236 | -.221 | -.072 | .164 | .225 | .236 | .000 | .359 | .281 |
| Information about other treatment options (midwife) | .065 | .057 | -.022 | -.155 | .157 | .181 | .093 | .641 | .211 |
| Encouraged to ask questions about diseases, treatment and care (midwife) | .105 | .010 | .131 | .006 | -.068 | .201 | -.059 | .694 | .233 |
| Information clearly explained (nurse) | .032 | -.139 | -.041 | .013 | .057 | .720 | .085 | .188 | .234 |
| Information about other treatment options (nurse) | .009 | .079 | -.062 | -.078 | .079 | .701 | .085 | .225 | .172 |
| Encouraged to ask questions about diseases, treatment and care (nurse) | .065 | .047 | .124 | -.036 | -.090 | .699 | -.045 | .227 | .216 |

Table 8.5 Promax rotated factor solution for three birth phase (continued)

| Factor Name | Confiden- tiality | Choice and continuity | Dignity | Prompt attention | Autonomy | Communi- cation | Quality of basic amenities | Social consi- deration | Unique variance |
|---|----------------------|-----------------------------|---------|---------------------|----------|--------------------|----------------------------------|------------------------------|--------------------|
| Experience of the waiting time when arriving on the place of delivery | .154 | .031 | -.102 | .743 | -.121 | .012 | -.051 | .005 | .423 |
| Experience of the waiting time on examinations | -.108 | .109 | .030 | .683 | .003 | -.001 | -.096 | .399 | .257 |
| Experience of the waiting time after you asked for help (midwife) | -.056 | .196 | .072 | .724 | -.068 | -.101 | -.031 | -.014 | .360 |
| Accessibility by phone (midwife) | -.026 | .013 | .188 | .082 | .089 | .099 | .270 | .080 | .493 |
| Travelling time to the place of birth | .082 | -.144 | .085 | .514 | .215 | .038 | .028 | -.174 | .463 |
| Experience of the waiting time after you asked for help (nurse) | .012 | .034 | .030 | .406 | -.023 | .118 | .234 | -.002 | .387 |
| Accessibility by phone (nurse) | .055 | .164 | -.020 | .195 | -.033 | .237 | .218 | -.023 | .402 |
| Facilitate the support of relatives and friends (midwife) | .144 | .025 | .000 | -.064 | -.098 | -.032 | .937 | -.017 | .253 |
| Consideration of home situation when planning appointments/examinations | .123 | .101 | .089 | -.121 | .190 | .018 | .466 | .119 | .253 |
| Facilitate the support of relatives and friends (nurse) | -.091 | .075 | .050 | .077 | -.057 | .219 | .690 | -.026 | .287 |
| Hygiene of the toilets and examination rooms. | -.068 | -.086 | .244 | .167 | .206 | .108 | .151 | -.160 | .503 |
| Comfort of the examination rooms and waiting rooms | .157 | -.002 | .256 | .303 | .201 | -.108 | .005 | -.241 | .440 |
| Quality of the food | -.078 | .237 | .143 | .171 | .193 | -.153 | -.039 | -.028 | .667 |
| Able to choose own health care provider (midwife) | .119 | .611 | -.116 | .124 | .032 | -.264 | .144 | .132 | .383 |
| Able to be referred to a medical specialist (midwife) | .267 | .398 | .127 | -.165 | .069 | -.062 | .014 | .074 | .404 |
| Presence of different health care providers (midwife) | .062 | .622 | -.039 | .056 | .066 | .128 | -.096 | .043 | .295 |
| Continuity of care by one health care provider (midwife) | .059 | .434 | .148 | .179 | .034 | -.003 | -.094 | .119 | .321 |
| Able to choose own health care provider (nurse) | -.207 | .683 | -.123 | .020 | .077 | -.062 | .344 | -.016 | .327 |
| Presence of different health care providers (nurse) | -.043 | .760 | -.002 | -.029 | -.076 | .495 | -.169 | -.150 | .269 |
| Continuity of care of one health care provider (nurse) | -.045 | .492 | -.024 | .049 | -.101 | .558 | .065 | -.131 | .294 |

Table 8.4 displays the transformed scores per domain and phase (1-10 scale). Mean transformed scores were positively skewed (7.1–8.4) as were the median scores (7.2–7.8). Floor effects were observed for up to 0.6% of women responding to a set of items in a particular domain for a particular phase, while ceiling effects were observed for up to 24% of cases. Mean scores and ceiling effects differed most across the domains in the antenatal phase and least across the domains in the post partum phase.

The Cronbach's alpha ranged from 0.57-0.82 for the antenatal phase, from 0.66-0.92 for the birth phase and from 0.62-0.95 for the post partum phase. For all phases the domain "quality of basic amenities" had lowest alphas.

Mean overall sum scores were higher for women whose child was not admitted after childbirth: 61.8 (SD 7.4) versus 58.3 (SD 5.1) ($p=0.02$) in the antenatal phase; 61.9 (SD 8.4) versus 57.9 (SD 7.7) ($p=0.06$) in the birth phase; and 62.1 (SD 9.2) versus 55.2 (SD 13.0) ($p=0.01$) in the post partum phase.

Eight factors corresponding to the domain structure of the generic responsiveness concept explained 69% of the variance in the antenatal phase, 69% in the birth phase and 76% in the post partum phase. Table 8.5 shows the final results of the oblique promax rotated factor loadings of the birth phase (the patterns of the antenatal and post partum phase were similar). Items that were expected to belong to one domain are bolded. The rotated solution of grouped items generally confirmed the hypothesized domain taxonomy within the birth and post partum phase. For the antenatal phase however, the hypothesized domain taxonomy was less evident with regard to "social consideration" and "choice and continuity", which appeared to be associated with other domains.

DISCUSSION

With the support of both patients and health care providers, we adapted the WHO's concept of responsiveness and the World Health Organization's responsiveness module into the ReproQ instrument to measure responsiveness in the Dutch obstetric care system antenatally, during childbirth and post partum. ReproQ was administered in a face-to-face interview context and appears to be a potential instrument for reporting perinatal service quality from the client's perspective. The perinatal responsiveness items grouped in the original 8-domain based structure found in the MCS Study and the World Health Survey questionnaire and appeared to be comprehensive, as judged by the experts. The ReproQ demonstrated satisfactory psychometric properties to describe the responsiveness outcomes of perinatal care in the Netherlands, with

preliminary evidence on the questionnaire's ability to discriminate between levels of non-clinical quality of care.

Particular strengths of adapting an existing WHO responsiveness concept and measurement approach are noted first. The 8 domains adopted a pre-existent conceptual structure that was identified in WHO's review of the patient satisfaction and quality of care literature, which also included the examination of different survey instruments.¹⁵ During this review, it was noted that the domain's value is supported by human rights law which argues that the responsiveness features of a health system are important in their own right.^{14,15,16} In contrast to patient satisfaction questionnaires, the concept of responsiveness tries to capture the patient's actual experience, since literature has shown that expectations strongly influence patient satisfaction. Expectations may be influenced by economic influences, political influences, prior experiences and socio-demographic characteristics.¹⁷⁻²⁰ Responsiveness aims to develop a universal concept which allows valid comparisons across different countries, ethnicities or health care systems.¹⁴ The responsiveness concept is challenged by a number of issues. Firstly, although responsiveness aims to measure the patient's actual experience, it is still disturbed by at least some extent of 'subjectivity'. Secondly, capturing responsiveness by a limited number of questions with fixed answering categories is quite challenging. Combining qualitative research and different (quantitative) survey techniques, one can produce a richer, more valid, and more reliable findings than when adopting qualitative or quantitative methods alone.²⁸

In spite of the existing strengths of the responsiveness concept and measurement approach, our study contributes to addressing some of the challenges. This includes whether it can truly be adapted to specific areas of health systems, like perinatal care, and elicit participation from specific groups of user interacting with specific health services. In particular, we found that users of perinatal services were interested in participating in the survey on non-clinical aspects of their care experience. Participation rates were equal or higher than the participation rates found in other perinatal satisfaction studies.^{29,30,31} Participation rates were equal to participation rates found in surveys measuring similar domains of quality of care, and better than obtained by WHO's Multi-Country Survey (MCS) study administered in the Netherlands in 2001 (59%).^{21,25,32} Comparisons are made with the MCS Study that was conducted in the Netherlands in 2001 as the questionnaire contained multiple items for each responsiveness domain, whereas the subsequent World Health Survey only contained one question per domain.^{14,21}

An optimal data collection method includes one with explicit trade-off balance between cost and errors including non-sampling error, coverage error, non-response error and measurement error.³³ To ensure data quality we chose face-to-face interviews with an independent interviewer for data collection. Compared to self-administered forms face-to-face interviews per-

form better in terms of non-sampling and non-response error but may perform worse when sensitive questions are asked and are more costly.³⁴ Internet or web surveys are less costly and more time efficient but also have limitations especially with coverage error.^{35,36} Mixed-mode approaches, combining the best of both worlds (being less costly and having less error than in a unimode approach) are very promising and should be considered.³⁷

The average item missing rates across domains was 8%, which according to literature can be considered acceptable. Within the framework of the MCS Study, a slightly lower overall missing rate was reported (5.0%).³⁸ However, our survey dealt with a group of women who were extremely occupied with the challenging demands of a new life, being interviewed post partum. Our survey focused on three phases of a specific health event, which may have been more cognitively demanding than the MCS Study, which focused on reporting on an average experience in the previous 12 months, and was shorter (on average 25 minutes).^{31,39,40} As found in the MCS, the domain missing rate was highest for the domains of “autonomy” and “choice and continuity” which are typically cognitively demanding domains. Across phases of perinatal care, the missing rate was highest for birth phase. But in general we found the proportion of missing rates per item to be similar across items. There were 2 items in the birth phase that had notably higher missing rates. Most likely this is the consequence of these items pointing to service events that do not always take place. For this paper we excluded them from the analysis. We do not feel this hindered our ability to test the ReproQ psychometric properties, but have noted difficulties with these items for future surveys. Alternatively, when not all women experience all the events that can occur, different responsiveness scores may be presented for certain service events that occurred as well as in absence of those events.

The transformed scale scores were satisfactory. A floor effect was almost absent as is frequently the case in positive-skewed assessments of self-reported health or self-rated experiences of (maternity) care.^{30,31,41,42} There was surprisingly less skewing towards use of the most positive category (ceiling effects) compared to other surveys.^{31,39}

Within each phase and for all domains, the questionnaire’s internal consistency was good. Cronbach’s alpha coefficients in ReproQ were similar compared to the CAHPS and WHO surveys^{21,25}, except for the domain “quality of basic amenities” which showed poor alphas in all phases. This domain contained questions about sanitary hygiene, comfort of waiting room and quality of food. It can be argued that these elements of basic amenities were too diverse to achieve internal consistency (see Table 8.5) and one might improve reporting of results from the questionnaire by analysing these items separately. The Pregnancy and Childbirth questionnaire (PCQ) covering personal treatment patient satisfaction outcomes for the antenatal and birth

phase⁴³ showed higher Cronbach's alpha coefficients for the antenatal phase (0.89 vs. 0.73) and for the birth phase (0.86 vs. 0.84). However, no predefined domain structure was used.

Overall, the taxonomy of domains from the WHO concept and measurement approach held for the adapted items in the ReproQ across all phases, although this taxonomy was weaker in the antenatal phase. This could possibly be due to factors such as; recall bias introduced by assessing all phases together, contamination by pregnancy outcome, focusing on one particular event or the heterogeneity in measurements since antenatal care consists of multiple visits. Underlying patterns are still to be explored. One may consider presenting a questionnaire on the antenatal phase within the antenatal phase, separately from a questionnaire on the birth and post partum phase. The total explained variance for the birth phase was higher in our study compared to the PCQ⁴³ (69% vs. 56%) as for the antenatal phase (69% vs. 53%).

The ability of the instrument to discriminate between good and less good experiences will be of paramount importance for its future use. We found some promising test results. The respondents clearly expressed different opinions on their experiences in the different phases of perinatal care. The non-uniform pattern of domain scores across the three phases suggested that respondents judged each phase separately as was intended by the questionnaire design. Furthermore discrimination between women whose infants were admitted to hospital subsequent to birth, was reflected in the lower sumscores across all phases. However, to test the difference in mean responsiveness of the birth phase between the mothers whose infant was hospitalised and the mothers whose infants were not hospitalised (mean difference: 3.8, pooled SD 6.5), at least 194 mothers had to be included in the analysis (type I error=0.05 (two-sided), power=0.80, control/case-ratio: 6/1). This implies that different responses on antenatal sumscores may reflect a true outcome on non-clinical aspect of care or may be contaminated by pregnancy outcomes. This again stresses the need to present a questionnaire on the antenatal phase separately.

Test-retest reliability was not performed in this stage. Reasons were to avoid the burden for the participants and to avoid associated potential recall bias effects due to having at this time a too demanding interview.

CONCLUSIONS

Overall, our study found that ReproQ demonstrated satisfactory psychometric properties to describe the responsiveness outcomes of perinatal care in the Netherlands, with preliminary evidence supporting the questionnaire's ability to discriminate between levels of non-clinical quality of care. In general, psychometric properties were in line with results obtained for other survey instruments that have been tested and promoted as part of quality assessment effort. In conclusion, given the overall favorable study results, we feel that this unique adaptation of the WHO responsiveness questionnaire to evaluate the various phases of perinatal care has been relatively successful. With some minor adaptations as suggested throughout the discussion we believe that this questionnaire can be used to evaluate the quality of perinatal care in the Netherlands and elsewhere.

REFERENCES

1. Ravelli AC, Tromp M, van Huis M, Steegers EA, Tamminga P, Eskes M, Bonsel GJ. Decreasing perinatal mortality in The Netherlands, 2000-2006: a record linkage study. *J Epidemiol Commun H.* 2009;63(9):761-5.
2. Campbell SM, Roland MO, Buetow SA. Defining quality of care. *Soc Sci Med.* 2000; 51(11):1611-25.
3. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med.* 1997;45(12):1829-43.
4. Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. *Soc Sci Med.* 1998;47(10):1573-88.
5. McPake B. User charges for health services in developing countries: a review of the economic literature. *Soc Sci Med.* 1993;36(11):1397-405.
6. Ware JE, Snyder MK, Wright WR, Davies AR. Defining and measuring patient satisfaction with medical care. *Eval Prog Plann.* 1983;6(3-4):247-93.
7. Wouters A. Essential national health research in developing countries: health care financing and the quality of care. *Int J Health Plan M.* 1991;6:253-71.
8. Gilson L, Allilio M, Heggenhougen K. Community satisfaction with primary health care services: an evaluation undertaken in the Morogoro region of Tanzania. *Soc Sci Med.* 1994;39(6):767-80.
9. Wensing M, Elwyn G. Methods for incorporating patients' views in health care. *Brit Med J.* 2003;326(7394):877-9.
10. Greener I. Towards a history of choice in UK health policy. *Social Health Ill.* 2009; 31(3):309-24.
11. Christiaens W, Gouwy A, Bracke P. Does a referral from home to hospital affect satisfaction with childbirth? A cross-national comparison. *BMC Health Serv Res.* 2007;7:e109. doi: 10.1186/1472-6963-7-109.
12. Wiegers TA. The quality of maternity care services as experienced by women in the Netherlands. *BMC Preg Childbirth.* 2009(May 9);9:e18. doi: 10.1186/1472-6963-12-69.
13. Wiegers TA, van der Zee J, Keirse MJ. Transfer from home to hospital: what is its effect on the experience of childbirth? *Soc Sci Med.* 1998;46(11):1505-11.
14. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB. Health system responsiveness: concepts, domains and measurement. In: Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism.* Geneva: World Health Organization; 2003. p. 573-96.
15. de Silva A. A framework for measuring responsiveness. *Global programme on evidence for health policy discussion paper no. 32.* Geneva: World Health Organization; 2000.
16. Gostin L, Hodge J, Valentine N, Nygren-Krug H. The domains of health responsiveness: a human rights analysis. *Health and human rights working paper no. 2.* Geneva: World Health Organization; 2003. [online: http://apps.who.int/iris/bitstream/10665/73926/1/HHRWPS2_eng.pdf]
17. Blendon RJ, Kim M, Benson JM. The public versus the World Health Organization on health system performance. *Health Affair.* 2001;20(3):10-20.

Chapter 8. Health System Responsiveness by N Valentine

18. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Satisfaction, gender, and communication in medical visits. *Med Care*. 1994;32(12):1216–31.
19. Murphy-Cullen CL, Larsen LC. Interaction between the socio-demographic variables of physicians and their patients: its impact upon patient satisfaction. *Soc Sci Med*. 1984;19(2):163–6.
20. Murray CJL, Kawabata K, Valentine NB. People's experience versus people's expectations. *Health Affair*. 2001;20(3):41–3.
21. Valentine NB, Bonsel GJ, Murray CJL. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Qual Life Res*. 2007;16(7):1107–25.
22. Health System Change. An update on the community tracking study: a focus on the changing health system. Issue Brief. Centre for Studying Health System Change, Mathematica Policy Research. Issue brief no 18. 1999(Feb). [online: www.hschange.org/content/63/63.pdf]
23. Cleary PD, Edgman-Levitan S, Walker JD, Gerteis M, Delbanco TL. Using patient reports to improve medical care: a preliminary report from 10 hospitals. *Qual Manag Health Care*. 1993;2(1):31–8.
24. Sixma HJ, van Campen C, Kerssens JJ, Peters L. Quality of care from the perspective of elderly people: the QUOTE-elderly instrument. *Age Aging*. 2000;29(2):173–8.
25. Hargraves JL, Hays RD, Cleary PD. Psychometric properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 adult core survey. *Health Serv Res*. 2003;38(6 Pt 1):1509–27.
26. Nunnally JCB, Bernstein IH. *Psychometric theory*. 3rd ed. New York: McGraw-Hill; 1994.
27. Gandek B, Ware JE, Jr., Aaronson NK, Alonso J, Apolone G, Bjorner J, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51(11):1149–58.
28. Frechtling J, Sharp L. *The user-friendly handbook for mixed-method evaluations*. Arlington, VA: Directorate for Education and Human Resources; 1997(August). Contract no.: 94-52965. Sponsored by the National Science Foundation. [online: www.nsf.gov/pubs/1997/nsf97153/start.htm]
29. Smith LF. Development of a multidimensional labour satisfaction questionnaire: dimensions, validity, and internal reliability. *Qual Health Care*. 2001;10(1):17–22.
30. Smith LF. Postnatal care: development of a psychometric multidimensional satisfaction questionnaire (the WOMBPNSQ) to assess women's views. *Br J Gen Pract* 2011, 61(591)e628–37.
31. Redshaw M, Heikkila K. *Delivered with care: a national survey of women's experience of maternity care*. Oxford: National Perinatal Epidemiology Unit; 2010. [online: <https://www.npeu.ox.ac.uk/delivered-with-care>]
32. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Health Care*. 2002;14(5):353–8.
33. Biemer PP, Lyberg LE. *Introduction to survey quality*. New York: John Wiley; 2003.
34. de Leeuw E. *Quality in mail, telephone, and face-to-face surveys*. Amsterdam: TT-Publicaties; 1992.
35. Couper M. Web surveys: a review of issues and approaches. *Public Opin Quart*. 2000; 64(4):464–94).
36. Dillman DA. *Mail and internet surveys: the tailored design method*. New York: John Wiley; 2000.

37. de Leeuw E. To mix or not to mix data collection modes in surveys. *J Off Stat.* 2005; 21(2):233–55.
38. Üstün TB, Chatterji S, Villanueva M, Bendib L, Celik C, Sadana R, et al. WHO Multi-Country Survey Study on Health and Responsiveness 2000-2001. In Murray CJL, Evans DB, editors. *Health systems performance assessment*. Geneva:World Health Organization; 2003. p. 761–96.
39. Vujkovic M, Birnie E, Bonsel GJ. *Maternal postnatal care utilisation in Rotterdam 2008-2009*. The Rotterdam Public Health Report. Rotterdam: GGD Rotterdam Rijnmond; 2011.
40. Demott K, Bick D, Norman R, Ritchie G, Turnbull N, Adams C, Barry C, Byrom S, Elliman D, Marchant S, Mccandlish R, Mellows H, Neale C, Parkar M, Tait P, Taylor C. *Clinical Guidelines and evidence review for post natal care: routine post natal care of recently delivered women and their babies*. London: National Collaborating Centre for Primary Care (NICE) and Royal College of General Practitioners; 2006. [online: <http://www.nice.org.uk/nicemedia/live/10988/30146/30146.pdf>]
41. Peterson WE, Charles C, DiCenso A, Sword W. The Newcastle satisfaction with nursing scales: a valid measure of maternal satisfaction with inpatient postpartum nursing care. *J Adv Nurs.* 2005;52(6):672–81.
42. Smith LF. Development of a multidimensional labour satisfaction questionnaire: dimensions, validity, and internal reliability. *Qual Health Care.* 2001;10(1):17–22.
43. Truijens SE, Pommer AM, van Runnard Heimel PJ, Verhoeven CJ, Oei SG, Pop VJ. Development of the Pregnancy and Childbirth Questionnaire (PCQ): evaluating quality of care as perceived by women who recently gave birth. *Eur J Obstet Gynecol Reprod Biol.* 174(March 2014):35–40.

Appendix 8.1 Missing values, mean(SD), range, percentage floor and ceiling given for each item separately

| Antenatal phase | No. of missing | % of missing | Mean | SD | Range | % Floor | % Ceiling | 25th %tile | 50th %tile | 75th %tile |
|------------------|----------------|--------------|------|------|-------|---------|-----------|------------|------------|------------|
| Dignity | 2 | 1.2% | 4.31 | 0.55 | 2.00 | 5.00 | 0.00 | 33.92 | 4.00 | 5.00 |
| | 3 | 1.8% | 4.20 | 0.62 | 2.00 | 5.00 | 0.00 | 29.24 | 4.00 | 5.00 |
| | 2 | 1.2% | 4.34 | 0.61 | 2.00 | 5.00 | 0.00 | 40.35 | 4.00 | 5.00 |
| Autonomy | 12 | 7.0% | 3.96 | .76 | 1.00 | 5.00 | .58 | 18.71 | 4.00 | 4.00 |
| | 19 | 11.1% | 3.92 | .69 | 2.00 | 5.00 | .00 | 14.62 | 3.95 | 4.00 |
| | 10 | 5.8% | 4.12 | .62 | 1.00 | 5.00 | .58 | 21.05 | 4.00 | 4.00 |
| Confidentiality | 4 | 2.3% | 4.11 | .65 | 1.00 | 5.00 | .58 | 22.81 | 4.00 | 4.00 |
| | 8 | 4.7% | 4.15 | .49 | 3.00 | 5.00 | .00 | 19.88 | 4.00 | 4.00 |
| | 12 | 7.0% | 4.11 | .57 | 1.00 | 5.00 | .58 | 19.30 | 4.00 | 4.00 |
| Communication | 3 | 1.8% | 4.06 | .74 | 1.00 | 5.00 | 1.17 | 23.98 | 4.00 | 4.00 |
| | 2 | 1.2% | 4.00 | .66 | 2.00 | 5.00 | .00 | 18.13 | 4.00 | 4.00 |
| | 11 | 6.4% | 3.81 | .77 | 1.00 | 5.00 | 1.17 | 12.87 | 3.21 | 4.00 |
| Prompt attention | 3 | 1.8% | 4.06 | .74 | 2.00 | 5.00 | .00 | 25.73 | 4.00 | 5.00 |
| | 3 | 1.8% | 3.96 | .67 | 1.00 | 5.00 | .58 | 16.37 | 4.00 | 4.00 |
| | 2 | 1.2% | 3.17 | 1.04 | 1.00 | 5.00 | 7.60 | 6.43 | 2.41 | 3.00 |

Appendix 8.1 continued

| | | No. of missing | % of missing | Mean | SD | Range | % Floor | % Ceiling | 25th %tile | 50th %tile | 75th %tile |
|-----------------------------------|--|----------------|--------------|------|------|-----------|---------|-----------|------------|------------|------------|
| | How did you experience the waiting time after you asked for help? | 15 | 8.8% | 3.81 | .82 | 1.00 5.00 | 1.17 | 14.62 | 3.00 | 4.00 | 4.00 |
| | How well was the accessibility by phone? | 2 | 1.2% | 3.79 | .94 | 1.00 5.00 | 3.51 | 17.54 | 4.00 | 4.00 | 4.00 |
| | How do you rate the travel time to your health service? | 2 | 1.2% | 4.12 | .72 | 1.00 5.00 | 1.75 | 25.73 | 4.00 | 4.00 | 5.00 |
| Social consideration | Did the health care provider facilitate the support of your relatives and friends? | 4 | 2.3% | 3.63 | .91 | 1.00 5.00 | 1.17 | 12.28 | 3.00 | 4.00 | 4.00 |
| | Was the home situation taken into consideration when planning an appointment? | 7 | 4.1% | 3.76 | .83 | 1.00 5.00 | 1.75 | 13.45 | 3.00 | 4.00 | 4.00 |
| Quality of basic amenities | How do you rate the quality of the hygiene of the toilets? | 3 | 1.8% | 3.98 | .77 | 1.00 5.00 | 1.75 | 19.30 | 4.00 | 4.00 | 4.00 |
| | How do you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanliness? | 3 | 1.8% | 3.78 | .75 | 2.00 5.00 | .00 | 13.45 | 3.00 | 4.00 | 4.00 |
| Choice and continuity | Were you able to choose your own health care provider? | 8 | 4.7% | 3.67 | 1.04 | 1.00 5.00 | 5.85 | 15.20 | 3.00 | 4.00 | 4.00 |
| | Were you able to use other health care services other than the one you usually went to? | 18 | 10.5% | 3.93 | .66 | 1.00 5.00 | .58 | 12.87 | 3.92 | 4.00 | 4.00 |
| | How well was the continuity of care by a single health care provider managed? | 4 | 2.3% | 3.69 | .97 | 1.00 5.00 | 4.68 | 14.62 | 3.00 | 4.00 | 4.00 |
| Birth phase | | | | | | | | | | | |
| Dignity | Respect shown during physical examinations (midwife) | 7 | 4.1% | 4.05 | .66 | 2.00 5.00 | .00 | 19.30 | 4.00 | 4.00 | 4.00 |
| | Examination room suitable to provide privacy | 9 | 5.3% | 4.14 | .61 | 2.00 5.00 | .00 | 23.98 | 4.00 | 4.00 | 5.00 |
| | Treated with respect (midwife) | 7 | 4.1% | 4.23 | .63 | 2.00 5.00 | .00 | 30.99 | 4.00 | 4.00 | 5.00 |
| | Respect shown during physical examinations (nurse) | 8 | 4.7% | 4.14 | .61 | 2.00 5.00 | .00 | 23.39 | 4.00 | 4.00 | 4.02 |
| | Treated with respect (nurse) | 8 | 4.7% | 4.19 | .70 | 2.00 5.00 | .00 | 30.99 | 4.00 | 4.00 | 5.00 |
| Autonomy | Involved in making a decision regarding your examinations or treatments | 18 | 10.5% | 4.01 | .68 | 2.00 5.00 | .00 | 16.96 | 4.00 | 4.00 | 4.00 |
| | Able to refuse examinations or treatments | 24 | 14.0% | 3.93 | .71 | 1.00 5.00 | .58 | 13.45 | 4.00 | 4.00 | 4.00 |
| | Asked permission before testing or starting treatment | 18 | 10.5% | 3.99 | .68 | 2.00 5.00 | .00 | 15.79 | 4.00 | 4.00 | 4.00 |

Appendix 8.1 continued

| | | No. of missing | % of missing | Mean | SD | Range | % Floor | % Ceiling | 25th %tile | 50th %tile | 75th %tile |
|------------------|--|----------------|--------------|------|-----|-----------|---------|-----------|------------|------------|------------|
| Confidentiality | Protecting your confidentiality during consultations (midwife) | 18 | 10.5% | 4.02 | .59 | 2.00 5.00 | .00 | 14.04 | 4.00 | 4.00 | 4.00 |
| | Confidentiality kept on provided information (midwife) | 15 | 8.8% | 4.06 | .55 | 2.00 5.00 | .00 | 16.37 | 4.00 | 4.00 | 4.00 |
| | Confidentiality of patients' medical records preserved (midwife) | 18 | 10.5% | 4.04 | .56 | 2.00 5.00 | .00 | 14.62 | 4.00 | 4.00 | 4.00 |
| | Protecting your confidentiality during consultations (nurse) | 17 | 9.9% | 3.91 | .50 | 1.00 5.00 | .58 | 14.62 | 4.00 | 4.00 | 4.00 |
| | Confidentiality kept on provided information (nurse) | 19 | 11.1% | 4.06 | .50 | 3.00 5.00 | .00 | 14.04 | 4.00 | 4.00 | 4.00 |
| | Confidentiality of patients' medical records preserved (nurse) | 19 | 11.1% | 4.05 | .57 | 2.00 5.00 | .00 | 15.79 | 4.00 | 4.00 | 4.00 |
| Communication | Information clearly explained (midwife) | 9 | 5.3% | 4.10 | .64 | 1.00 5.00 | .58 | 21.64 | 4.00 | 4.00 | 4.00 |
| | Information about other treatment options (midwife) | 19 | 11.1% | 3.95 | .68 | 1.00 5.00 | .58 | 13.45 | 4.00 | 4.00 | 4.00 |
| | Encouraged to ask questions about diseases, treatment and care (midwife) | 8 | 4.7% | 4.05 | .66 | 1.00 5.00 | .58 | 19.30 | 4.00 | 4.00 | 4.00 |
| | Information clearly explained (nurse) | 8 | 4.7% | 4.05 | .66 | 1.00 5.00 | .58 | 19.88 | 4.00 | 4.00 | 4.00 |
| | Information about other treatment options (nurse) | 20 | 11.7% | 3.93 | .69 | 1.00 5.00 | .58 | 13.45 | 4.00 | 4.00 | 4.00 |
| | Encouraged to ask questions about diseases, treatment and care (nurse) | 7 | 4.1% | 4.02 | .70 | 1.00 5.00 | .58 | 19.30 | 4.00 | 4.00 | 4.00 |
| Prompt attention | Experience of the waiting time when arriving on the place of delivery | 7 | 4.1% | 3.98 | .96 | 1.00 5.00 | 1.75 | 30.41 | 4.00 | 4.00 | 5.00 |
| | Experience of the waiting time on examinations | 11 | 6.4% | 3.88 | .91 | 1.00 5.00 | 2.92 | 93.57 | 4.00 | 4.00 | 4.00 |
| | Experience of the waiting time after you asked for help (midwife) | 12 | 7.0% | 3.89 | .91 | 1.00 5.00 | 1.17 | 22.81 | 3.00 | 4.00 | 4.00 |
| | Accessibility by phone (midwife) | 31 | 18.1% | 4.07 | .68 | 2.00 5.00 | .00 | 19.88 | 4.00 | 4.00 | 4.00 |
| | Travelling time to the place of birth | 10 | 5.8% | 4.15 | .69 | 1.00 5.00 | 1.17 | 26.32 | 4.00 | 4.00 | 5.00 |
| | Experience of the waiting time after you asked for help (nurse) | 21 | 12.3% | 3.88 | .79 | 2.00 5.00 | .00 | 16.96 | 3.59 | 4.00 | 4.00 |
| | Accessibility by phone (nurse) | 29 | 17.0% | 4.00 | .72 | 1.00 5.00 | 1.17 | 16.37 | 4.00 | 4.00 | 4.00 |

Appendix 8.1 continued

| | No. of missing | % of missing | Mean | SD | Range | % Floor | % Ceiling | 25th %tile | 50th %tile | 75th %tile |
|---|----------------|--------------|------|-----|-----------|---------|-----------|------------|------------|------------|
| Social consideration | | | | | | | | | | |
| Facilitate the support of relatives and friends (midwife) | 15 | 8.8% | 3.89 | .81 | 1.00 5.00 | .58 | 16.96 | 4.00 | 4.00 | 4.00 |
| Consideration of home situation when planning appointments/examinations | 39 | 22.8% | 3.89 | .75 | 1.00 5.00 | .58 | 13.45 | 3.83 | 4.00 | 4.00 |
| Facilitate the support of relatives and friends (nurse) | 15 | 8.8% | 3.95 | .79 | 1.00 5.00 | .58 | 19.30 | 4.00 | 4.00 | 4.00 |
| Quality of basic amenities | | | | | | | | | | |
| Hygiene of the toilets and examination rooms. | 9 | 5.3% | 4.13 | .72 | 2.00 5.00 | .00 | 28.07 | 4.00 | 4.00 | 5.00 |
| Comfort of the examination rooms and waiting rooms | 14 | 8.2% | 4.06 | .74 | 1.00 5.00 | .58 | 23.98 | 4.00 | 4.00 | 5.00 |
| Quality of the food | 19 | 11.1% | 3.61 | .89 | 1.00 5.00 | 2.34 | 11.11 | 3.00 | 4.00 | 4.00 |
| Choice and continuity | | | | | | | | | | |
| Able to choose own health care provider (midwife) | 24 | 14.0% | 3.52 | .97 | 1.00 5.00 | 2.3 | 8.77 | 3.00 | 4.00 | 4.00 |
| Able to be referred to a medical specialist (midwife) | 54 | 31.6% | 3.94 | .73 | 1.00 5.00 | 1.17 | 11.11 | 4.00 | 4.00 | 4.00 |
| Presence of different health care providers (midwife) | 14 | 8.2% | 3.82 | .86 | 1.00 5.00 | 1.17 | 14.62 | 4.00 | 4.00 | 4.00 |
| Continuity of care by one health care provider (midwife) | 9 | 5.3% | 3.85 | .95 | 1.00 5.00 | 2.34 | 21.05 | 4.00 | 4.00 | 4.00 |
| Able to choose own health care provider (nurse) | 24 | 14.0% | 3.48 | .99 | 1.00 5.00 | 2.92 | 9.36 | 3.00 | 4.00 | 4.00 |
| Presence of different health care providers (nurse) | 14 | 8.2% | 3.78 | .81 | 1.00 5.00 | 1.75 | 10.53 | 4.00 | 4.00 | 4.00 |
| Continuity of care of one health care provider (nurse) | 11 | 6.4% | 3.89 | .82 | 1.00 5.00 | 1.75 | 15.79 | 4.00 | 4.00 | 4.00 |
| Post Partum phase | | | | | | | | | | |
| Dignity | | | | | | | | | | |
| Respect shown during physical examinations (midwife) | 11 | 6.4% | 4.08 | .60 | 2.00 5.00 | .00 | 18.71 | 4.00 | 4.00 | 4.00 |
| Examination room suitable to provide privacy | 18 | 10.5% | 4.02 | .81 | 1.00 5.00 | 2.34 | 20.47 | 4.00 | 4.00 | 4.00 |
| Treated with respect (midwife) | 10 | 5.8% | 4.14 | .60 | 2.00 5.00 | .00 | 22.81 | 4.00 | 4.00 | 4.06 |
| Respect shown during physical examinations (nurse) | 8 | 4.7% | 4.07 | .70 | 1.00 5.00 | 1.17 | 19.88 | 4.00 | 4.00 | 4.00 |
| Treated with respect (nurse) | 7 | 4.1% | 4.09 | .79 | 1.00 5.00 | 1.75 | 25.73 | 4.00 | 4.00 | 5.00 |
| Autonomy | | | | | | | | | | |
| Involved in making a decision regarding your examinations or treatments | 17 | 9.9% | 3.90 | .83 | 1.00 5.00 | 1.17 | 16.37 | 4.00 | 4.00 | 4.00 |
| Able to refuse examinations or treatments | 24 | 14.0% | 3.88 | .84 | 1.00 5.00 | 1.75 | 15.20 | 4.00 | 4.00 | 4.00 |
| Consider your needs related to pregnancy and birth | 19 | 11.1% | 3.97 | .74 | 1.00 5.00 | .58 | 17.54 | 4.00 | 4.00 | 4.00 |

Appendix 8.1 continued

| | | No. of missing | % of missing | Mean | SD | Range | % Floor | % Ceiling | 25th %tile | 50th %tile | 75th %tile |
|-----------------------------|--|----------------|--------------|------|-----|-----------|---------|-----------|------------|------------|------------|
| | Help you to decide about the treatment of pain during childbirth | 20 | 11.7% | 3.84 | .88 | 1.00 5.00 | .6 | 15.20 | 4.00 | 4.00 | 4.00 |
| | Decided about the place of birth | 27 | 15.8% | 3.86 | .87 | 1.00 5.00 | 2.92 | 13.45 | 4.00 | 4.00 | 4.00 |
| Confidentiality | Protecting your confidentiality during consultations (midwife) | 12 | 7.0% | 3.97 | .74 | 1.00 5.00 | 3.51 | 18.13 | 4.00 | 4.00 | 4.00 |
| | Confidentiality kept on provided information (midwife) | 17 | 9.9% | 4.05 | .67 | 1.00 5.00 | .58 | 16.96 | 4.00 | 4.00 | 4.00 |
| | Confidentiality of patients' medical records preserved (midwife) | 20 | 11.7% | 4.02 | .66 | 1.00 5.00 | 1.17 | 15.79 | 4.00 | 4.00 | 4.00 |
| | Protecting your confidentiality during consultations (nurse) | 11 | 6.4% | 3.98 | .72 | 1.00 5.00 | .58 | 16.96 | 4.00 | 4.00 | 4.00 |
| | Confidentiality kept on provided information (nurse) | 18 | 10.5% | 3.99 | .68 | 1.00 5.00 | .58 | 14.04 | 4.00 | 4.00 | 4.00 |
| | Confidentiality of patients' medical records preserved (nurse) | 18 | 10.5% | 3.96 | .69 | 1.00 5.00 | 1.17 | 14.04 | 4.00 | 4.00 | 4.00 |
| Communication | Information clearly explained (midwife) | 7 | 4.1% | 3.85 | .88 | 1.00 5.00 | 1.17 | 21.05 | 4.00 | 4.00 | 4.00 |
| | Information about other treatment options (midwife) | 25 | 14.6% | 3.81 | .90 | 1.00 5.00 | 1.75 | 14.62 | 4.00 | 4.00 | 4.00 |
| | Encouraged to ask questions about diseases, treatment and care (midwife) | 8 | 4.7% | 3.98 | .82 | 1.00 5.00 | 2.34 | 19.88 | 4.00 | 4.00 | 4.00 |
| | Information clearly explained (nurse) | 8 | 4.7% | 3.95 | .85 | 1.00 5.00 | 2.34 | 21.64 | 4.00 | 4.00 | 4.00 |
| | Information about other treatment options (nurse) | 18 | 10.5% | 3.88 | .87 | 1.00 5.00 | 2.34 | 17.54 | 4.00 | 4.00 | 4.00 |
| | Encouraged to ask questions about diseases, treatment and care (nurse) | 9 | 5.3% | 4.00 | .79 | 1.00 5.00 | 1.75 | 20.47 | 4.00 | 4.00 | 4.00 |
| Prompt attention | Health care available if urgent | 9 | 5.3% | 4.03 | .87 | 1.00 5.00 | 1.75 | 27.49 | 4.00 | 4.00 | 5.00 |
| | Health care available if not urgent | 26 | 15.2% | 3.93 | .89 | 1.00 5.00 | 2.34 | 19.30 | 4.00 | 4.00 | 4.00 |
| | Health care services were accessible | 13 | 7.6% | 3.96 | .82 | 1.00 5.00 | 2.34 | 18.71 | 4.00 | 4.00 | 4.00 |
| | Caregivers were reachable by phone or a call to the bed | 12 | 7.0% | 3.95 | .88 | 1.00 5.00 | 2.34 | 21.64 | 4.00 | 4.00 | 4.00 |
| Social consideration | Did your caregivers also focus on your family and household? | 14 | 8.2% | 3.93 | .86 | 1.00 5.00 | 1.75 | 19.88 | 4.00 | 4.00 | 4.00 |

Appendix 8.1 continued

| | No. of missing | % of missing | Mean | SD | Range | % Floor | % Ceiling | 25th %tile | 50th %tile | 75th %tile |
|---|----------------|--------------|------|-----|-----------|---------|-----------|------------|------------|------------|
| Consideration of home situation when planning appointments/examinations | 21 | 12.3% | 3.91 | .81 | 1.00 5.00 | 2.34 | 15.20 | 4.00 | 4.00 | 4.00 |
| Facilitate the support of relatives and friends (nurse) | 10 | 5.8% | 3.95 | .95 | 1.00 5.00 | 4.68 | 22.81 | 4.00 | 4.00 | 4.14 |
| Quality of basic amenities | | | | | | | | | | |
| Cleanliness of the surroundings | 4 | 2.3% | 3.94 | .86 | 1.00 5.00 | 1.75 | 21.64 | 4.00 | 4.00 | 4.00 |
| Comfort of the bed | 8 | 4.7% | 3.79 | .86 | 1.00 5.00 | 2.34 | 16.37 | 3.00 | 4.00 | 4.00 |
| Choice of meals | 9 | 5.3% | 3.80 | .94 | 1.00 5.00 | 3.51 | 19.88 | 3.00 | 4.00 | 4.00 |
| Choice and continuity | | | | | | | | | | |
| Making service time available on request of the client | 25 | 14.6% | 3.57 | .99 | 1.00 5.00 | 2.92 | 11.11 | 3.00 | 4.00 | 4.00 |
| Continuity of care provision when change of individual (within profession) | 18 | 10.5% | 3.75 | .92 | 1.00 6.00 | 3.51 | 11.11 | 4.00 | 4.00 | 4.00 |
| Continuity of care provision when change of professional (across professions) | 21 | 12.3% | 3.15 | .80 | 1.00 5.00 | 2.92 | 9.36 | 3.00 | 4.00 | 4.00 |
| Allowance for selecting a preferred type of health care provider | 13 | 7.6% | 3.83 | .89 | 1.00 5.00 | 3.51 | 14.62 | 4.00 | 4.00 | 4.00 |
| Being explicit on which health professional is actual in charge | 9 | 5.3% | 3.90 | .81 | 1.00 5.00 | 1.75 | 18.13 | 4.00 | 4.00 | 4.00 |

CHAPTER 9

Quality of perinatal care services from the user's perspective: a Dutch study applies the World Health Organization's responsiveness concept



J. van der Kooy, E. Birnie, N.B. Valentine, J.P. de Graaf, S. Denktas,
E.A.P. Steegers, G.J. Bonsel

BMC Pregnancy and Childbirth. 2017; 17(327): 1-11

ABSTRACT

Background. The concept of responsiveness was introduced by the World Health Organization (WHO) to address non-clinical aspects of service quality in an internationally comparable way. Responsiveness is defined as aspects of the way individuals are treated and the environment in which they are treated during health system interactions. The aim of this study is to assess responsiveness outcomes, their importance and factors influencing responsiveness outcomes during the antenatal and delivery phases of perinatal care.

Methods. The Responsiveness in Perinatal and Obstetric Health Care Questionnaire was developed in 2009/10 based on the 8-domain WHO concept and the World Health Survey questionnaire. After ethical approval, a total of 171 women, who were 2 weeks postpartum, were recruited from three primary care midwifery practices in Rotterdam, the Netherlands, using face-to-face interviews. We dichotomized the original 5 ordinal response categories for responsiveness attainment as 'poor' and 'good' responsiveness and analyzed the ranking of the domain performance and importance according to frequency scores. We used a series of independent variables related to health services and users' personal background characteristics in multiple logistic regression analyses to explain responsiveness.

Results. Poor responsiveness outcomes ranged from 5.9% to 31.7% for the antenatal phase and from 9.7% to 27.1% for the delivery phase. Overall for both phases, "respect for persons" (autonomy, dignity, communication and confidentiality) domains performed better and were judged to be more important than "client orientation" domains (choice and continuity, prompt attention, quality of basic amenities, social consideration). On the whole, responsiveness was explained more by health-care and health related issues than personal characteristics.

Conclusions. To improve responsiveness outcomes caregivers should focus on domains in the category "client orientation".

BACKGROUND

The performance of perinatal care is often judged by endpoints such as perinatal morbidity, mortality and costs. However, quality of care literature supports the view that non-clinical aspects of health care, such as service quality, are important aspects of the system's performance too and, moreover, may affect clinical outcomes.¹⁻³ Better service quality is thought to increase compliance with medical treatment, and to improve information transfer and utilization of health services.⁴⁻⁷ Governments of Western countries increasingly acknowledge the importance of incorporating non-clinical service quality when the performance of the system is monitored.^{8,9}

An important approach to measuring service quality is the concept of “responsiveness”, which was introduced by the World Health Organization in the World Health Report 2000 to compare service quality in an internationally comparable way. Responsiveness is defined as aspects of the way individuals are treated and the environment in which they are treated during health system interactions.¹⁰ Aspects refer to non-financial, non-clinical qualities of care that reflect respect for human dignity and interpersonal aspects of the care process. Being based on utility theory, the concept separates the utility individuals derive from clinical and non-clinical aspects, and from a policy perspective can be used to make trade-offs between non-clinical quality and clinical quality. Utility theory refers to the measurement of preferences over some set of goods and services.¹⁰ Human rights law argues that the responsiveness features of a health system are important in their own right.¹⁰⁻¹²

Perinatal care in the Netherlands is organized as a system of inter-related services, that include referral practices, covering the different phases of the perinatal experience: antenatal care, delivery and postpartum care. Perinatal care is provided by independently operating community midwives providing care for low-risk pregnant women (primary healthcare) and obstetricians and gynaecologists providing in-hospital care for high-risk women (secondary and tertiary care). Most women receive postpartum care by a community midwife. Most perinatal deaths occur during the antenatal and delivery phases.¹³ International studies and the National Report of the Netherlands reported that the perinatal mortality rate in 2004 for the Netherlands was the highest in Europe (10.5/1,000 live births). In 2010 the perinatal mortality rate declined (9.0/1,000 live births).¹⁴⁻¹⁷ As a result, the evaluation of the different aspects of perinatal care, in particular the antenatal and delivery phases, is crucial.¹⁸ Evaluation of non-clinical aspects of the quality of care may be even more important, since the majority of women are not ill. This may increase the importance of non-clinical aspects. Thus far, few attempts have been made to evaluate non-clinical quality of the perinatal health care system across the different phases. The few studies available observe that aspects of health care services influence patient satisfaction.¹⁹⁻²¹ They did not investigate the different phases of the perinatal system nor use an internationally

comparable questionnaire. To our knowledge, ours is the first study to present this information for perinatal care in the Netherlands. The aim of our study was to assess the responsiveness outcomes and factors influencing responsiveness outcomes of perinatal health care in urban settings in the Netherlands using the Responsiveness in Perinatal and Obstetric Health Care Questionnaire, the ReproQ questionnaire, which was based on the WHO concept of health system responsiveness and modified from existing WHO questionnaires.

METHODS

Questionnaire

The Responsiveness in Perinatal and Obstetric Health Care Questionnaire (ReproQ) was developed between October 2009 and February 2010. The ReproQ is based on the same 8 domains identified for measuring responsiveness in WHO's review of the patient satisfaction and quality of care literature. The 8 domains were dignity, autonomy, confidentiality, communication (collectively categorized as the "respect for persons" domains), choice and continuity, prompt attention, quality of basic amenities, and social consideration (collectively categorized as "client orientation" domains). To build the ReproQ, slight adaptations were made to the set of responsiveness questions translated from the WHO questionnaires.^{10,22}

The ReproQ was designed, as with most of the WHO questionnaires, to be administered in a face-to-face interview setting. The ReproQ asks essentially the same set of questions for the three different phases of perinatal care but, for purposes of this paper, we focus on two phases - the antenatal and delivery phases - the most important for the infant mortality challenge mentioned earlier. More importantly, postpartum care different in its characteristics and delivery site since it is delivered only at home and includes only home nurses and midwives limiting discussion of referral practices. In addition it includes evaluation of paediatric care. Data on the postnatal care will therefore be studied separately. The antenatal phase was defined as the period from the onset of pregnancy until the onset of delivery. Respondents were asked to provide an overall evaluation of their experiences that took place during the antenatal care period rather than just a single visit that may be biased in either way by a particular incident. The delivery phase referred to the period of birth.

Each phase was covered by the above mentioned 8 domains, with 2-7 question items per domain. The standardized response options consisted of five verbal response categories: "very good", "good", "moderate", "bad", and "very bad". In total, 65 responsiveness question items were distributed over two phases (25 antenatal, 40 delivery). In addition, 29 question items on personal and healthcare-related characteristics associated with the experience were also included.

Table 9.1 shows the 8 domains and question items for the antenatal phase. The domains and items for the delivery phase were roughly similar. The following psychometric properties of the ReproQ were evaluated: feasibility, reliability and validity. Feasibility: the interviews lasted between 20 and 40 minutes and the overall missing rate was 8%. Construct validity: mean Cronbach's alphas for the antenatal, birth and postpartum phase were: 0.73 (range 0.57-0.82), 0.84 (range 0.66-0.92), and 0.87 (range 0.62-0.95) respectively. The item-own scale correlations within all phases were considerably higher than most of the item-other scale correlations. Within the antenatal, birth and postpartum phase, the 8 factors explained 69%, 69%, and 76% of variance respectively. Discriminative validity: overall responsiveness mean sum scores were higher for women whose children were not admitted, as expected from literature. The ReproQ interview-based questionnaire demonstrated satisfactory psychometric properties, with the potential to discriminate between quality of care levels. Detailed descriptions of the reported psychometric properties are reported elsewhere.²³

Study population; data collection

The study was a cross-sectional, interview-based survey of women having had a delivery in the previous 2 weeks. Study approval was granted by the Medical Ethical Committee, Erasmus Medical Centre, Rotterdam, the Netherlands, no MEC2012207. Study respondents were recruited from three primary care midwifery practices in the urban area of Rotterdam, the Netherlands, between February 2010 and March 2011. These three practices were geographically chosen since they provide care for almost all women living at the north side of Rotterdam. Within these three midwife practices 25 different community midwives provide care. Women or their partners or family were required to speak and understand Dutch sufficiently, the latter serving as translators rather than proxy respondents. Written informed consent was obtained prior to the interview. Study interviews were carried out by 10 trained and independent interviewers, but first invitations to participate in the study were made by the respondents' own midwife at the postpartum visit 2 weeks after delivery. Respondents were invited in a consecutive order, using the day of delivery. The interview was usually held at the home of the respondent. Interviewees were invited to respond to all questions, yet never forced to. The average interview was 30 minutes long.

Table 9.1 The 8 domains with the (REPROQ) question items formulated for the antenatal phase

| Respect for persons | |
|---|--|
| Autonomy | <p>How well were you involved in making decisions regarding your examinations or treatments?</p> <p>Were you able to refuse examinations or treatments?</p> <p>Were you asked permission before testing or starting treatment?</p> |
| Dignity | <p>Were physical examinations and treatments done in a way that respected your privacy?</p> <p>Did the examination rooms ensure your privacy?</p> <p>Were you treated with respect by your health care provider?</p> |
| Communication | <p>How well were things explained by your health care provider in a way you could understand?</p> <p>Was written information provided in such a way you could understand?</p> <p>Were you encouraged to ask questions about your health problems, treatment and care?</p> <p>Were you given time to ask questions about your health problem or treatment?</p> <p>Was information on the health service's contact, location and parking information clear to you?</p> |
| Confidentiality (of information) | <p>Were consultations carried out in a manner that protected your confidentiality?</p> <p>Was confidentiality kept on the information provided by you?</p> <p>Was your medical record kept confidential?</p> |
| Client orientation | |
| Choice and continuity (of health care provider) | <p>Were you able to choose your own health care provider?</p> <p>Were you able to use other health care services other than the one you usually went to?</p> <p>How well was the continuity of care by a single provider facilitated?</p> <p>Were you able to choose your own place of delivery?</p> |
| Prompt attention | <p>How well did you receive prompt attention at your health service?</p> <p>How did you experience the waiting time after you asked for help?</p> <p>How do you rate the accessibility of the health care provider by phone?</p> <p>How do you rate the travel time to your health service?</p> |
| Quality of basic amenities | <p>How do you rate the quality of the hygiene of the toilets?</p> <p>How do you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanness?</p> <p>How do you rate the quality of the food?</p> |
| Social consideration | <p>Did the health care provider facilitate the support of your relatives and friends?</p> <p>Was the home situation taken into consideration when planning an appointment?</p> |

Responsiveness measures and background characteristics

Two responsiveness outcome measures were estimated to describe performance: question and domain measures. For question measures, the five options answers were grouped into binary categories; 'good' and 'poor'. The 'poor' rating was used when a respondent reported the item as either "very bad", "bad" or "moderate". For domain measures, if over 33% of the items were rated poor within a domain, the rating of 'poor' was used for the whole domain. The percentage approach was used to score domains as the number of question items per domain differed across domains. Dichotomization was chosen as it has been shown to reduce bias caused by reporting scale contraction for disadvantaged groups. Relevant differences in non-optimal outcomes can therefore be missed. For similar reasons and for reasons of the right-skewedness of the data, we chose to judge a domain as poor when at least 33% of the items were judged as poor.^{24, 25} Dichotomization avoids problems associated with violating regression assumptions when testing which personal or health service factors are associated with responsiveness. Another commonly used metrics, sum scores, were positively skewed but we chose not to use this metric due to it being less useful for addressing reporting behaviour. Thresholds were selected ex ante and results were presented in the same way the WHO Responsiveness reports were presented in the past (http://www.who.int/responsiveness/papers/MCSS_Analytical_Guidelines.pdf).

Responsiveness domain importance measures were calculated based on individual rankings of the set of domains.

Background characteristics with a feasible influence over the responsiveness performance rating were chosen. These were: parity (nulliparous/multiparous), age (≤ 30 / >30 years), ethnicity (Dutch/non-Dutch), education level (low or middle/high), marital status (single/relationship or married), living in a deprived neighborhood (yes/no, based on 4-digit zip-codes and a public list of deprived zip-code based neighbourhoods issued by the Dutch government)²⁶, Dutch language proficiency (good/weak or poor), obstetric history (yes/no, based on self-report of mother or child outcomes which required a medical intervention by a gynaecologist), adverse child outcome (yes/no, based on self-reported asphyxia, (possible) congenital anomaly, infection, small for gestational age (child too small), and/or premature birth), paediatric hospital admission within 1 month (yes/no), receiving pain medication when requested (yes/no), receiving an intervention (yes/no, instrumental delivery or a caesarean section), maternal hospital admission during the antenatal period or within 1 month after birth (yes/no), day of delivery (weekend/weekday), time of delivery (8-18 h /18-8 h), healthcare pathway during pregnancy (referral to secondary care during antenatal or birth care, yes/no), perinatal healthcare pathway (start antenatal care with midwife, not referred; start antenatal care with midwife, referred during antenatal care to gynaecologist; start antenatal care with midwife, referred during birth care to gynaecologist; antenatal and birth care with gynaecologist).

Data handling

Records of a respondent were regarded missing if all scores on all phases were missing (antenatal, delivery and postpartum phase). If response was partial, the response was evaluated per phase. Respondents were excluded for one phase if all items were missing for that phase. Values for missing question items were imputed with the mean when only up to 3 items were missing in a particular phase. We imputed these values to increase precision and power. We imputed with the mean as a conservative approach. Bias is toward non-significance, hereby not overestimating associations in our results.^{27,28} Variables with over 30% missing values were not imputed and excluded from analysis.

We assumed a baseline proportion of poor performance per domain of 10% (i.e. 90% of respondents has sufficient score); we further assumed that the difference between non-referred patients during delivery, and referred patients was substantial and relevant, expecting from several sources doubling i.e. 20% poor performers (see e.g., Rijnders 2008 (29)). Under these assumptions $n=196$ is sufficient to discriminate between these known groups. Higher rates in general (as we actually observed in half of the domains) lower the sample size, smaller difference increase the size. Our final sample size was 171 respondents, implying that our analysis was slightly underpowered for the assumptions.

Analyses

Descriptive analyses

Data were analysed using SPSS software version 17.0. Responsiveness performance and importance outcomes were described in frequency tables and by spider diagrams by phase. The assigned importance of each domain was plotted against domain scores (% good responsiveness) and visually inspected for any observed relationship.

Bivariate analyses

Spider diagrams were also used to show patterns in responsiveness outcomes between advantaged and disadvantaged subpopulations as was done in the WHO guide. (30) For these comparisons we grouped a set of disadvantaged subpopulations according to the following background characteristics: for "respect for persons" than for "client orientation" domains. In the antenatal phase, (1) multiparous, (2) Dutch-origin, (3) having started with a midwife, and not being referred, and (4) having no child hospitalization. The unpaired Student's t-test or the Chi square test were used to compare groups on these characteristics.

Regression analyses

Multiple regression was used to explain responsiveness for each domain by the background characteristics (personal and healthcare-related).

Forward stepwise analysis was used (inclusion $p < 0.05$; exclusion $p > 0.05$) to explain domain outcomes.

RESULTS

A total of consecutive 274 respondents were invited for participation, 180 respondents (66%) agreed to be interviewed. Reasons for non-participation included the anticipated time burden, feeling at unease having a stranger visit their home, and logistic reasons such as incorrect phone number, or incorrect address. Of the 180 interviews planned, seven interviews (4%) were cancelled by the women and two interviews (1%) were cut short because the respondent's language proficiency was inadequate and no translator was present. The remaining 171 interviews (95%) were analysed. Eighteen (11%) of these interviews were conducted with either translation or in English.

Table 9.2 describes the respondent's background characteristics. About 70% of women were between the ages of 25 and 34, only 4% had no or low education, half were of Dutch origin and about half came from underprivileged neighbourhoods, and most had a high proficiency of spoken Dutch. Related to health-care characteristics, 60% were primiparous. On the perinatal health care pathway, about a third started antenatal care with a midwife and were not referred, and another third started antenatal care with midwife, and were referred during birth care to a gynaecologist.

Table 9.3 shows the responsiveness performance outcome measures by domain for the antenatal and the delivery phases. The proportion of poor responsiveness outcomes ranged from 5.9% (dignity) to 31.7% (social consideration) in the antenatal phase and from 9.7% (dignity) to 27.1% (choice and continuity) in the delivery phase. For both phases, "respect for persons" (autonomy, communication, confidentiality, dignity) domains were judged better than the "client orientation" domains.

Table 9.2 Respondent's personal, health and health-care characteristics (n=171)

| Variable | Number | Percent |
|--|--------|---------|
| Personal characteristics | | |
| Maternal Age ^a | | |
| <19 years | 3 | 2% |
| 20-25 years | 15 | 9% |
| 25-34 years (REF) | 119 | 70% |
| >35 years | 33 | 19% |
| Missing | 1 | 1% |
| Parity | | |
| Primiparous | 97 | 57% |
| Multiparous (REF) | 74 | 43% |
| Education | | |
| Low | 6 | 4% |
| Middle | 75 | 44% |
| High (REF) | 90 | 53% |
| Marital status | | |
| Single | 30 | 18% |
| Relationship/married (REF) | 141 | 82% |
| Ethnic background | | |
| Dutch (REF) | 94 | 55% |
| Non Dutch | 77 | 45% |
| Neighbourhood | | |
| Privileged neighbourhood (REF) | 84 | 49% |
| Underprivileged neighbourhood | 87 | 51% |
| Proficiency (speaking) Dutch | | |
| Good/excellent (REF) | 153 | 89% |
| Weak/poor | 18 | 11% |
| Health-care characteristics | | |
| Obstetric history ^b | | |
| Primiparous | 97 | 57% |
| Multiparous, no medical history (REF) | 24 | 14% |
| Multiparous, medical history | 50 | 29% |
| Perinatal health care pathway | | |
| (1) Start antenatal care with midwife, not referred (REF) | 61 | 36% |
| (2) Start antenatal care with midwife, referred during antenatal care to gynaecologist | 37 | 22% |
| (3) Start antenatal care with midwife, referred during birth care to gynaecologist | 57 | 33% |
| (4) Antenatal and birth care with gynaecologist | 16 | 9% |

Table 9.2 Respondent's personal, health and health-care characteristics (continued)

| Variable | Number | Percent |
|---|--------|---------|
| Pain medication during labour | | |
| No request (REF) | 79 | 46% |
| No pain medication received after requesting | 32 | 19% |
| Pain medication received after requesting | 58 | 34% |
| Intervention during labour^c | | |
| No (REF) | 97 | 57% |
| Yes, no emergency intervention | 51 | 30% |
| Yes, emergency intervention | 21 | 12% |
| Day of delivery | | |
| Weekend | 37 | 22% |
| Weekday (REF) | 134 | 78% |
| Time of delivery | | |
| 0-8hr | 45 | 26% |
| 8-18hr (REF) | 82 | 48% |
| 18-24hr | 43 | 25% |
| Missing | 1 | 1% |
| Adverse outcome of child^d | | |
| No adverse outcome (REF) | 128 | 75% |
| Adverse outcome | 43 | 25% |
| Hospital admission of child | | |
| No admission (REF) | 145 | 85% |
| Admission | 26 | 15% |
| Hospital admission of the mother | | |
| No admission (REF) | 154 | 90% |
| Admission | 17 | 10% |

Key: REF=reference in logistic regression; ^a Mean age 30 (range 18-42); ^b Obstetric history based on self reported mother or child outcomes which required intervention of a gynaecologist; ^c Caesarean section or instrumental delivery; ^d Adverse outcome based on self reported asphyxia (shortage of oxygen), (possible) congenital anomaly, infection, small for gestational age (child too small), premature birth

Table 9.3 Client reported poor responsiveness for each domain, for the antenatal and delivery phase separately

| Domain | Antenatal Phase | | Delivery Phase | |
|---------------------------------|------------------------|--|------------------------|--|
| | Number of participants | Percentage reporting poor responsiveness | Number of participants | Percentage reporting poor responsiveness |
| Respect for persons | | | | |
| Autonomy (AU) | 161 | 18.0% | 155 | 15.7% |
| Dignity (DI) | 169 | 5.9% | 165 | 9.7% |
| Communication (CM) | 168 | 20.0% | 166 | 14.2% |
| Confidentiality (CF) | 159 | 7.8% | 153 | 11.6% |
| Client orientation | | | | |
| Choice and continuity (CC) | 167 | 28.1% | 162 | 27.1% |
| Prompt attention (PA) | 169 | 30.0% | 144 | 20.6% |
| Quality of basic amenities (QA) | 168 | 22.9% | 156 | 23.4% |
| Social consideration (SC) | 164 | 31.7% | 158 | 22.1% |

Domain importance measures were higher (higher frequency) , for the “respect for persons” domains than the “client orientation” domains (average 69%; 95% CI 60%-76% versus 31%; 95%CI 24%-40%). The highest importance was assigned to the domains of communication (26%) and dignity (22%) and the lowest was assigned to choice and continuity (6%) and social consideration (4%). Of similar importance were autonomy, confidentiality, prompt attention, and quality of basic amenities (range: 10%-11%).

Figure 9.1 (and Appendix 9.1) compares responsiveness performance between subpopulations. In all disadvantaged (including more at risk) subpopulations the proportion of poor responsiveness was lower. Multiparous women tended to show poorer responsiveness outcomes on nearly all domains. The same pattern was found in women with an obstetric history. Ethnic differences were mainly observed within the antenatal phase where Dutch women showed poorer responsiveness outcomes. Women living in a deprived neighbourhood and those who did not speak Dutch proficiently tended to have the same responsiveness pattern. Groupings by neighbourhood showed no marked differences for the antenatal phase, while the delivery phase had marked difference in responsiveness outcomes. The nature of the differences in patterns between subgroups are mainly observed for the “client orientation” domains.

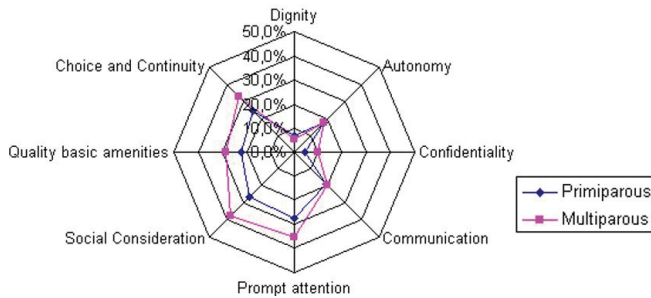
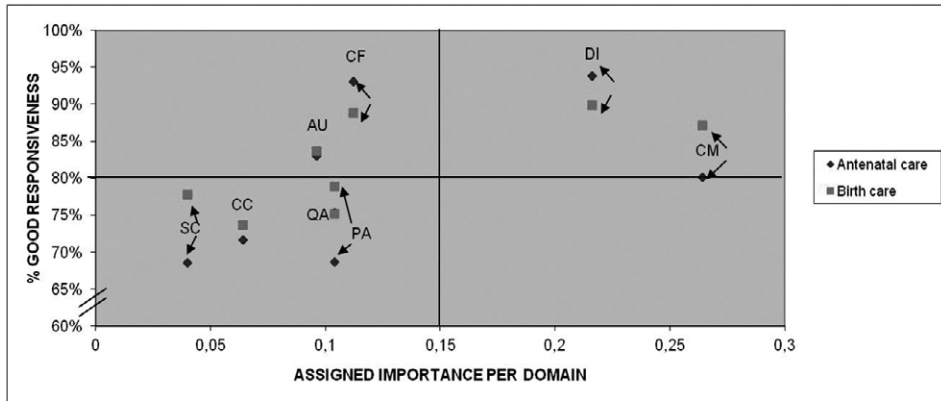


Figure 9.1 A comparison of the pattern of responsiveness quality for antenatal and birth phases by parity



Autonomy (AU), Dignity (DI), Confidentiality (CF), Communication (CM), Choice and Continuity (CC), Prompt Attention (PA), Quality of Basic Amenities (QA), Social Consideration (SC)

Figure 9.2 Comparison of the importance assigned to the responsiveness domains and the performance of domains: antenatal and birth phases

The relationship between the proportion of good domain performance and importance was roughly linear as seen in Figure 9.2. Of the domains of high importance, average good performance in the communication domain was above 80% (only 20% rating poor) but better in delivery versus antenatal phases. Of the domains of medium importance, prompt attention had low performance (more than 20% rating poor responsiveness), and performance differed widely between antenatal (poorer responsiveness) and birth phases.

Table 9.4 shows the odds ratios from the multiple logistic regression analyses for responsiveness outcomes and background characteristics, by domain, stratified by phase. Overall, health-care (service) characteristics were stronger predictors of responsiveness outcome performance than users' personal characteristics. In particular, health service characteristics were strong predictors in "client orientation" domain regressions for delivery and birth phases. No background characteristic was significantly associated with "respect for person" domains in the birth phase.

Specific domain-background characteristics associations for the antenatal and birth phases were as follows. Higher odds of prompt attention problems in both antenatal and birth phases was associated with obstetric history, and having an intervention. Hospital admission of the child was significant in the birth phase only. For choice and continuity and social consideration in the antenatal phase, having a non-Dutch background (ethnicity) was associated with lower odds of responsiveness problems (OR range: 0.27-0.42). In the birth phase, for choice and continuity, only respondents with parity had significantly reduced odds of problems (OR 0.25). Whereas, for social consideration obstetric history and hospital admission were associated with higher odds of responsiveness problems (OR range: 2.44, 3.23). For quality of basic amenities, only increased maternal age was significantly associated with higher odds of poor responsiveness in the antenatal phase.

DISCUSSION

Responsiveness quality of perinatal health services in the Netherlands, was better for "respect for persons" domains compared with "client orientation" domains. These are also domains that have more importance to users. Overall, the health status and health-care related characteristics of users explained more of the variation in responsiveness quality than personal characteristics (e.g., education, deprived neighbourhood) in the birth phase, while in the antenatal phase responsiveness is more associated with personal background characteristics.

We observed poorer responsiveness outcomes for the "client orientation" domains, than for the "respect for persons" domains. Similar results were found by Liabsuetrakul et al. and Qing Luo et al.^{31,32} This might be, because the domains of autonomy, dignity, communication and confidentiality are easier to change. They could be influenced by professionals changing behaviour over short periods of time instead of changes in the organization of care required for "client orientation" domains, which requires management coordination and longer time periods to implement changes. A second explanation might be that the domains in this category are judged as more important by the health professionals and thus are given more attention.

Table 9.4 Variance of reported poor outcome given for each domain for both the antenatal and birth phase. Only Odds Ratio's (95%CI) for significant determinants are given^a

| Domain | Antenatal Phase | | | | | Delivery Phase | | | | |
|----------------------------|-------------------|------|-------|------|---------|---|--------------|--------------|--------------|----------------|
| | Determinants | OR | 95%CI | | p-value | Determinants | OR | 95%CI | | p-value |
| Respect for persons | | | | | | Respect for persons | | | | |
| Autonomy | Intervention | 3.00 | 1.44 | 6.26 | 0.003 | None | | | | |
| Dignity | None | | | | | None | | | | |
| Communication | None | | | | | None | | | | |
| Confidentiality | Parity | 0.33 | 0.12 | 0.87 | 0.025 | None | | | | |
| Client orientation | | | | | | Client orientation | | | | |
| Choice and continuity | Ethnic background | 0.39 | 0.20 | 0.79 | 0.008 | Parity | 0.25 | 0.10 | 0.62 | 0.003 |
| Prompt attention | Obstetric history | 2.34 | 1.08 | 5.04 | 0.030 | Obstetric history | 4.11 | 1.54 | 10.99 | 0.005 |
| | Intervention | 2.42 | 1.14 | 5.11 | 0.021 | Hospital admission Child Intervention | 3.21 2.98 | 1.09 1.17 | 9.49 7.59 | 0.035 0.022 |
| Quality of basic amenities | Maternal age | 2.10 | 1.05 | 4.19 | 0.036 | None | | | | |
| Social consideration | Parity | 0.42 | 0.20 | 0.86 | 0.018 | Obstetric history | 2.44 | 1.05 | 5.67 | 0.038 |
| | Ethnic background | 0.27 | 0.13 | 0.57 | 0.001 | Hospital admission Child | 3.23 | 1.22 | 8.54 | 0.018 |

^a Inclusion $p < 0.05$; exclusion $p < 0.05$

On the whole, there was more variation in responsiveness explained by health-care and health related issues. Obstetric history and an adverse events (receiving an intervention) influenced responsiveness outcomes in the antenatal and birth phases. Other associated personal characteristics were also more health related – maternal age, parity. This is partially in line with reports from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) patient experience survey, which also showed the effect of health-related characteristics. These reports

observed more association between age, general health, education, individual health plan, and less association of ethnicity, gender and time in insurance plan with responses on patient experience.³³ Although we did not assess the impact of health plan and length of time in insurance plan, we observed similar associations to these for the importance of health and age (since age and parity are inversely related).

Other studies that assessed patient personal characteristics on (some) of the WHO responsiveness domains showed similar tendencies for the characteristics of parity, education and marital status. However these studies did not include birth outcomes within their analysis.^{31,34-36} One could only speculate to what extent differences are explained by cultural factors. More research on this area is needed.

Referral is a common feature of health care systems, in particular with the field of perinatal care. Being referred during pregnancy does not seem to impact responsiveness. This is in line with other studies, which found no association with being referred and responsiveness domains.^{19,21} However, some studies do find a negative association with being referred and patient satisfaction.²⁹

The domains of communication and dignity were most frequently identified as most important. This is partly in contrast with the population based survey conducted by WHO³⁷ and results by Liabsuetrakul et al., who assessed the importance of responsiveness domains in Thailand.³¹ They both found prompt attention and dignity to be the most important domains, followed by communication in third place. The preference for prompt attention in these other studies may be due to the fact that it was operationalized in terms of geographical access and access in case of emergencies. In our study prompt attention focussed upon waiting times. Results from other studies which also focussed upon waiting times support our results that prompt attention was valued as less important.^{2,3} Qing Luo et al. observed the domains basic amenities, communication and autonomy to be most important in community health services in China. They reasoned that prompt attention was well achieved and therefore not chosen as most important.³² Bramesfeld et al. saw a similar ranking as ours, but observed a difference in ranking between in- and outpatient mental care. Hereby, observing prompt attention to be more important in outpatient care.³⁸

The overall linear relationship we observed between good domain performance and assigned importance by users may be explained by health care professionals also judging these domains as important, and therefore placing more emphasis on them.

Our study had several strengths. Firstly, 66% of the invited women agreed to participate in this study. This is an effective study sample, since a response rate of 30% has been proposed as reasonable for patient satisfaction surveys and a response rate of 50% is considered to be quite high.^{13,14}

Secondly, our study covered many subpopulations in Rotterdam, including subpopulations which are often missed in satisfaction surveys. More frequent among non-participants in satisfaction studies are those having a language barrier, a psychiatric history, a low social economic status, a low educational level, no paid work and Muslim people.^{39,40} Since our study covered these subpopulations, its generalizability to women in perinatal care is more assured. Thirdly, interviews were conducted in such a way that known factors influencing respondent's health responsiveness outcomes were diminished as much as possible. Interviews were performed by independent interviewers, respondents were interviewed at their own homes and at an appropriate interval with respect to their birth experience (2 weeks postpartum). Previous studies have shown that women who answer surveys at home are more critical compared with respondents who are interviewed in the hospital, since the latter are loyal to the institution.⁴¹ Women being interviewed after 2 weeks also tend to be more critical.⁴²

A few limitations merit discussion. Firstly, since only people from urban areas participated in this study, the study population is presumably representative for Dutch urban areas, but the generalizability to the whole Dutch perinatal population remains uncertain. Secondly, translation could only be arranged for some of the women who did not understand the Dutch language sufficiently, this was done by a family member of the women. This could introduce a translation bias since this was not done by a professional translator. Thirdly, all non-Dutch ethnic groups were grouped resulting in a heterogeneous subpopulation. Responsiveness outcomes in these subpopulation may differ, since other studies showed that ethnicity can be of influence.²² Fourthly, no analysis was performed on non-participants. Fifthly, recall bias and carry over effects on health responsiveness outcomes within the antenatal phase cannot be excluded, since birth outcome determinants significantly influenced outcomes within the antenatal phase. Sixthly, we collected medical outcomes from the respondents themselves (self-report), which could lead to less accurate outcomes. In the future one could consider linking the survey data to the medical records. Lastly, the study was slightly underpowered given the power calculation.

Overall, our ReproQ questionnaire, which was directly derived from the WHO concept of responsiveness, was able to measure responsiveness outcomes of the perinatal care system in the Netherlands. As carry over effects on health responsiveness outcomes within the antenatal phase cannot be excluded we recommend that when evaluating the responsiveness outcomes of the perinatal health care system, antenatal care should be evaluated before the start of de-

livery to prevent carry over effects of birth outcomes. To improve responsiveness quality of the Dutch Perinatal Care system, caregivers should focus on domains covering the category “client orientation”.

REFERENCES

1. Campbell SM, Roland MO, Buetow SA. Defining quality of care. *Soc Sci Med.* 2000; 51(11):1611–25.
2. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med.* 1997;45(12):1829–43.
3. Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. *Soc Sci Med.* 1998;47(10):1573–88.
4. McPake B. User charges for health services in developing countries: a review of the economic literature. *Social science & medicine.* 1993;36(11):1397–405.
5. Ware JE, Snyder MK, Wright WR, Davies AR. Defining and measuring patient satisfaction with medical care. *Eval Prog Plann.* 1983;6(3-4):247–59.
6. Wouters A. Essential national health research in developing countries: health care financing and the quality of care. *Int J Health Plan M.* 1991;6:253–71.
7. Gilson L, Alilio M, Heggenhougen K. Community satisfaction with primary health care services: an evaluation undertaken in the Morogoro region of Tanzania. *Soc Sci Med.* 1994;39(6):767–80.
8. Greener I. Towards a history of choice in UK health policy. *Social Health Ill.* 2009; 31(3):309–24.
9. Wensing M, Elwyn G. Methods for incorporating patients' views in health care. *Brit Med J.* 2003;326(7394):877–9.
10. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB. Health system responsiveness: concepts, domains and measurement. In: Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism.* Geneva: World Health Organization; 2003. p. 573–96.
11. de Silva A. A framework for measuring responsiveness. Global programme on evidence for health policy discussion paper no. 32. Geneva: World Health Organization; 2000.
12. Gostin L, Hodge J, Valentine N, Nygren-Krug H. The domains of health responsiveness: a human rights analysis. Health and human rights working paper no. 2. Geneva: World Health Organization; 2003. [online: http://apps.who.int/iris/bitstream/10665/73926/1/HHRWPS2_eng.pdf]
13. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010. Rotterdam: Erasmus MC; 2010. [online: <http://www.nvk.nl/Nieuws/Dossiers/DossierPerinatalezorg.aspx>]
14. Mohangoo AD, Buitendijk SE, Hukkelhoven CW, Ravelli AC, Rijninks-van Driel GC, Tamminga P, Nijhuis JG. [Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study]. *Ned Tijdschr Geneesk.* 2008;152(50):2718–27.
15. Ravelli AC, Tromp M, van Huis M, Steegers EA, Tamminga P, Eskes M, Bonsel GJ. Decreasing perinatal mortality in The Netherlands, 2000–2006: a record linkage study. *J Epidemiol Commun H.* 2009;63(9):761–5.
16. Zeitlin J, Wildman K, Breart G, Alexander S, Barros H, Blondel B, Buitendijk S, Gissler M, Macfarlane A. PERISTAT: indicators for monitoring and evaluating perinatal health in Europe. *Eur J Public Health.* 2003;13(3 Suppl):29–37.

17. EURO-PERISTAT project in collaboration with SCPE EaE, 2008. European perinatal health report. Better Statistics for Better Health for Pregnant Women and Their Babies in 2004.
18. van der Velden J. Dutch: Een goed begin. Veilige zorg rond zwangerschap en geboorte [Internet]. Royal Dutch Organisation of Midwives (KNOV); 2009 [cited 2015 May 20]. Available from: http://www.knov.nl/docs/uploads/Advies_stg_zw_en_geb_Een_goed_beginpdf_spread-versie_voor_media.pdf
19. Christiaens W, Gouwy A, Bracke P. Does a referral from home to hospital affect satisfaction with childbirth? A cross-national comparison. *BMC Health Serv Res*. 2007;7:e109. doi: 10.1186.
20. Wiegers TA. The quality of maternity care services as experienced by women in the Netherlands. *BMC Preg Childbirth*. 2009(May 9);9:e18. doi: 10.1186/1472-6963-12-69.
21. Wiegers TA, van der Zee J, Keirse MJ. Transfer from home to hospital: what is its effect on the experience of childbirth? *Birth*. 1998;25(1):19–24.
22. Valentine NB, Bonsel GJ, Murray CJL. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Qual Life Res*. 2007;16(7):1107–25.
23. van der Kooy J, Valentine NB, Birnie E, Vujkovic M, de Graaf JP, Denktas S, Steegers EA, Bonsel GJ. Validity of a questionnaire measuring the World Health Organization concept of health system responsiveness with respect to perinatal services in the Dutch obstetric care system. *BMC Health Serv Res*. 2014;14(1):e622. doi: 10.1186/s12913-04-0622-1.
24. Valentine NB, Verdes-Tennant E, Bonsel GJ. Health systems' responsiveness and reporting behaviour: multilevel analysis of the influence of individual-level factors in 64 countries. *Soc Sci Med*. 2015;138:152–60.
25. Hosseinpoor AR, Stewart Williams JA, Itani L, Chatterji S. Socioeconomic inequality in domains of health: results from the World Health Surveys. *BMC Public Health*. 2012; 12:198.
26. VROM. Stad en Stijging. Sociale stijging als leidraad voor stedelijke vernieuwing. [City and Rise. Social advancement as a guideline for urban renewal]. The Hague. Publication advice No. 054. 2006.
27. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials*. 2004;1(4):368–76.
28. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol*. 2006;59(10):1087–91.
29. Rijnders M, Baston H, Schonbeck Y, van der Pal K, Prins M, Green J, Buitendijk S. Perinatal factors related to negative or positive recall of birth experience in women 3 years postpartum in the Netherlands. *Birth*. 2008;35(2):107–16.
30. World Health Organization. The health systems responsiveness analytical guidelines for surveys in the Multi-Country Survey Study 2005 [Internet]. Geneva: World Health Organization; 2005 (December) [cited 15 June 2016]. Available from: http://www.who.int/responsiveness/papers/MCSS_Analytical_Guidelines.pdf?ua=1
31. Liabsuetrakul T, Petmanee P, Sanguanchua S, Oumudee N. Health system responsiveness for delivery care in Southern Thailand. *Int J Qual Health Care*. 2012;24(2):169–75.
32. Luo Q, Wang Q, Lu Z, Liu J. Evaluation of responsiveness of community health services in urban China: a quantitative study in Wuhan City. *PLoS One*. 8(5):e62923.

33. Zweifler J, Hughes S, Lopez RA. Controlling for race/ethnicity: a comparison of California commercial health plans CAHPS scores to NCBD benchmarks. *Int J Equity Health*. 2010(Jan);9:e4. doi: 10.1186/1475-9276-9-4.
34. Raleigh VS, Hussey D, Seccombe I, Hallt K. Ethnic and social inequalities in women's experience of maternity care in England: results of a national survey. *J R Soc Med*. 2010;103(5):188–98.
35. Waldenstrom U, Nilsson CA, Winbladh B. The Stockholm birth centre trial: maternal and infant outcome. *Br J Obstet Gynaecol*. 1997;104(4):410–8.
36. Waldenstrom U, Rudman A, Hildingsson I. Intrapartum and postpartum care in Sweden: women's opinions and risk factors for not being satisfied. *Acta Obstet Gynecol Scand*. 2006;85(5):551–60.
37. Valentine N, Darby C, Bonsel GJ. Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of "health systems responsiveness" in 41 countries. *Social science & medicine*. 2008;66(9):1939–50.
38. Bramesfeld A, Wedegartner F, Elgeti H, Bisson S. How does mental health care perform in respect to service users' expectations? Evaluating inpatient and outpatient care in Germany with the WHO responsiveness concept. *BMC Health Serv Res*. 2007;7:e99.
39. Gayet-Ageron A, Agoritsas T, Schiesari L, Kolly V, Perneger TV. Barriers to participation in a patient satisfaction survey: who are we missing? *PLoS One*. 2011;6(10):e26852.
40. Boshuizen H, Viet A, Picavet H, Botterweck A, van Loon A. Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. *Public Health*. 2006;20(4):297–308.
41. Lumley J. Assessing satisfaction with childbirth. *Birth*. 1985;12(3):141–5.
42. van Teijlingen ER, Hundley V, Rennie AM, Graham W, Fitzmaurice A. Maternity satisfaction studies and their limitations: "What is, must still be best". *Birth*. 2003; 30(2):75–82.

Appendix 9.1 Spidergraphs showing responsiveness performance for different subpopulations

% of respondents with poor responsiveness outcomes given if the birth interval is primiparous

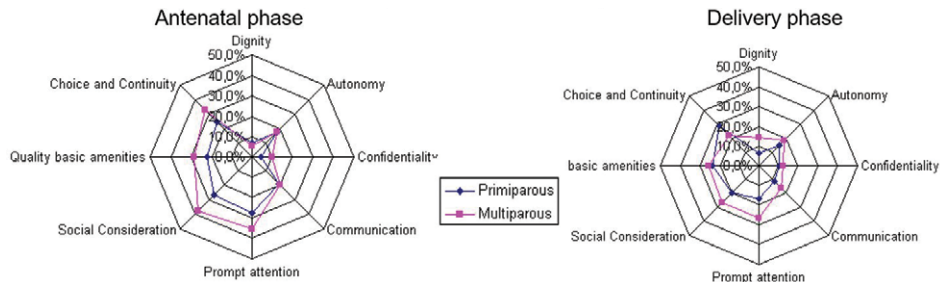


Figure A Comparison by parity

% of respondents with poor responsiveness outcomes given for ethnicity

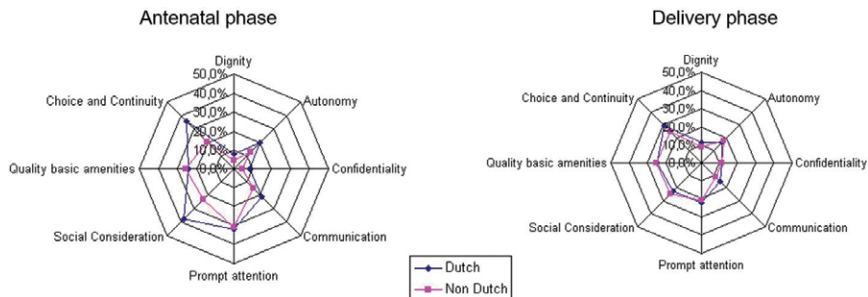


Figure B Comparison of responsiveness by ethnicity

% of respondents with poor responsiveness outcomes depending on the perinatal health care pathways

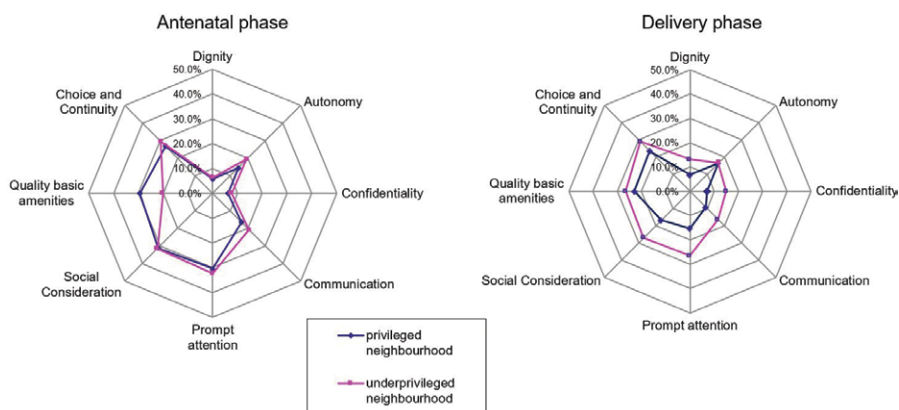


Figure C Comparison by privilege of neighbourhood

% of respondents with poor responsiveness outcomes depending on the perinatal health care pathways

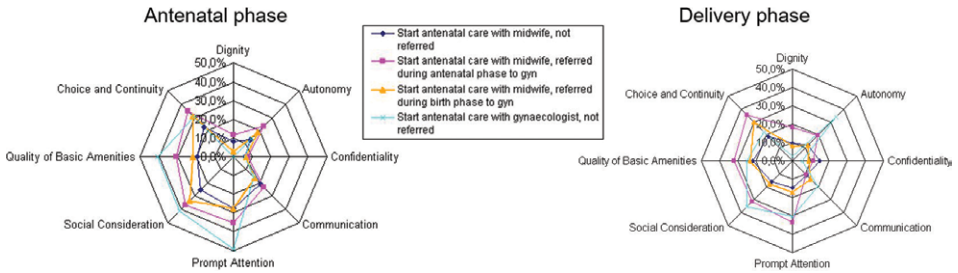


Figure D Comparison by perinatal health care pathway

% of respondents with poor responsiveness outcomes depending on hospital admission of the child

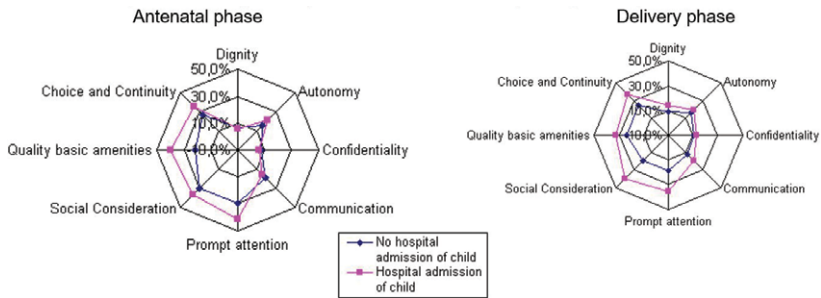


Figure E Comparison by hospital admission of the child

CHAPTER 10

General discussion



AIM OF THE THESIS

The main aim for this thesis was to enlarge the evidence base for the rationale and measurement of the World Health Organization's responsiveness concept. The initial proposition of the responsiveness concept was simple: to measure at the individual level of a client, user or patient an outcome covering service aspects of health care. This implied a radical change with Donabedian's 3-tier approach¹ with different levels of measurement. Service aspects - strictly thought of as personal experiences of the client with the health system, rather than as solely isolated procedural or structural features - were conceived as comprising 8 domains: "respect for persons" or 'personal' domains: autonomy, communication, confidentiality of information and dignity; and "client orientation" or 'setting' domains: choice, prompt attention, quality of basic amenities, and social support.

Universality of this concept across countries, cultures, health care systems, and within the health care system, apart from being of scientific interest, is important for global health. The initial proposal for the application of responsiveness was simple too: for comparison of health systems at the national level. This thesis empirically showed supportive evidence in favour of the responsiveness concept's universality, and showed its successful application in more settings than anticipated.

The claim for universality of the concept rests *inter alia* on sufficient universality of its measurement. The large scale WHO surveys represented a rich data source to explore cross-country and cross-person comparability. Standardized measurement instruments covering responsiveness were used as part of the WHO interview-administered surveys, covering 106 WHO surveys with approximately 258,000 respondents, and the 171 respondents to the specifically designed ReproQ survey in the The Netherlands added an important component of perinatal care.

Patient and to some extent country characteristics are a potential sources of bias. Through both traditional psychometric methods and innovative analyses this thesis supports the inter-individual and inter-country comparability of responsiveness measurement, after small adjustments required for harmonizing reporting scales. In other words: apart from medical outcomes like mortality and morbidity measures, a client-based service quality concept provided useful metrics.

Assuming the validity and universality of its measurement, the thesis investigated the relevance of the responsiveness concept and data to policy-makers and health service organizations. Indeed the responsiveness measures can be used for comparative health systems assessments in evaluations of global initiatives like 'universal health coverage'; even more, it could be translated

to evaluations of specific national clinical sub-systems, as illustrated with perinatal health care in the Netherlands. Perinatal health care services are an important test case given their significance across health systems, their cultural variation, and the acknowledged role of inequalities.²

This general discussion of the thesis key findings is organized as follows: **principle findings** according to the thesis questions; **cross-cutting themes** connected to the universality of the responsiveness concept; cross-cutting **methodological issues**; and a concluding set of **recommendations**.

PRINCIPAL FINDINGS

Is responsiveness a measurable and consistent concept for service quality, with universal meaning?

Psychometric measures of feasibility, reliability (temporal), and validity of both the Multi-Country Survey Study (2000-2001) and the World Health Survey (2002-2004) responsiveness questionnaires, appeared moderate to good (**chapters 2, 3**). The World Health Survey responsiveness questions made slight improvements on the MCS questions (**chapter 3**). Internal validity tests generally confirmed the hypothesized domain taxonomy of responsiveness. This adds to existing validity evidence provided by the global ordering of the vignettes.³ The switch in responsiveness, from measuring features of a system (the Donabedian approach) to counting quality experiences, paves the way for uniformly measuring patients/clients experiences. While conceptual universality is supported, its measurement is susceptible to reporting behaviour biases. These are impacts of characteristics of the respondent and her/his context (measured as individual-level characteristics), as well as her/his country characteristics on the way s/he uses the presented response options ('response scale'). Scale use is more affected by individual-level factors than hitherto assumed and showed some interesting domain-country interactions: individual use of the scale is particularly independent of country setting in 2 out of 8 domains (dignity and social support) but strongly dependent on country in 3 out of 8 domains (prompt attention, quality of basic amenities and confidentiality). International comparative analysis should account for these individual-level and country-level effects.

Importantly, the thesis succeeded in determining the relative importance of a large set of individual-level characteristics for explaining reporting behaviour, based on a theory of expectations. Under socio-demographic factors, education has the strongest effect; under health-related experiences, perceived health state, followed by caring for others with a chronic illness; and under health values, the assigned importance to responsiveness influenced the scale use. These new variables have larger effects than commonly used variables, namely age and sex. A further technical achievement was the subsequent demonstration of the relevance of taking account of

these effects. Of particular importance for analysis of health (care) equity is that lower educated populations avoid response scale end-points, and in particular, do not criticize poor performance; populations that are exposed to ill-health through caring for close others have stronger negative reactions to poor services, all other factors being equal, while those in ill-health themselves have pessimistic shifts with stronger negative reactions to better services. These findings imply that response bias adjustment should be considered during within-country comparative assessments with mixed populations (**chapter 4**).

There are several practical challenges to including reporting behaviour questions (vignettes) as part of an interview; a main one being that it takes time. But efficiency can be enhanced if one takes advantage of the observation that reporting behaviour is quite general across domains and items.

In conclusion, the responsiveness concept shows universality, and its uniform measurement is possible. To achieve optimal comparability a module of vignettes should be inserted in data collection for responsiveness, allowing for correction of reporting behaviour bias.

Having demonstrated support for the general concept and measurement: which responsiveness aspects matter most, if any?

The thesis conducted an investigation of responses on the importance of responsiveness domains using data from 42 surveys (including ReproQ) across Western and Eastern Europe, the Americas, and Asia. Taking the evidence across populations in different countries, the following 3 domains are the most important: communication; dignity; and prompt attention. The primacy of the dignity and communication domains was supported in the two general population surveys (the MCS, WHS) (**chapter 5**) and in the perinatal survey (**chapter 9**). The lowest endorsement was given to the domain of social support. Occasionally, geographical regions showed some smaller divergences (**chapters 3, 5**). Overall, this global uniformity of domain rankings contributes to construct validity of the responsiveness concept, and supports the comparative use of concept and data obtained.

Are there features of health systems and persons which - across countries - give better or poorer responsiveness?

This question is about true experiences, and the true personal and country background effects that are to be revealed and not adjusted for. The broad set of health service determinants (12), measured at the individual level, added to the personal background data (8), enabled a unique analysis covering 120,000 respondents from 49 countries (**chapter 6**). Perhaps the most striking result is that certain service characteristics are universally more important than personal characteristics, namely: the behaviour of health workers (discriminatory attitude externally

measured); actual (measured) travel time; and the service being managed by the government (rather than non-governmental).

For personal characteristics, rather than age and sex which have been commonly used in the literature, poverty and health status matter. These (and other) observations show responsiveness to be an equity-sensitive outcome for which specific pathways may depend on the domain studied. Services are not equal to all: we already knew this was the case when observing systematic variation in health outcomes according to social class, for example, but it is also true for service quality, which is a possible conduit for poor health outcomes. A simple example illustrates the point: being poor increases the presence of comorbidity⁴, which in turn increases the difficulty for health services to manage the care processes. Similarly, physical (rather than financial) access to health care facilities is reduced if means of transport are limited. In other words, poorer health erodes responsiveness, even ignoring financial mechanisms⁵, and reinforcing the so-called Inverse Care Law with the global observations made through this thesis: those in more need do not receive more services to the commensurate degree.

Aggregate ecological analyses investigated health system performance at the country level using variables which were defined at that level (e.g., national coverage rates for vaccination, antenatal visits etc.). The analyses assumed responsiveness to be instrumental for access and health outcomes. Essentially considering the instrumental value of responsiveness in this way is a more conventional, Donabedian-like approach. Controlling for population age and health service resources (an external variable in WHO's health systems assessment framework for responsiveness), better aggregate health system responsiveness was associated with higher aggregate intervention coverage (**chapters 3, 7**) and better aggregate health outcomes (**chapter 7**). These key associations at the aggregate level (using averages per country) converges with individual-level analysis (**chapter 6**), reducing the likelihood of an ecological fallacy. Specific responsiveness domains were sensitive to country-level features of the system or its health results. Dignity and prompt attention were significant in regressions for maternal mortality and TB mortality; dignity was significant in models predicting aggregate child mortality. Here the causal path is perhaps bidirectional. Whereas sometimes unidirectionality seems obvious. For example, the striking cross-country inequality in measles vaccination seems explained by both financial protection and responsiveness, indicating the relevance of 'acceptability' and 'affordability' quality^{5,6} for securing more intervention equality between the wealthy and poor (**chapter 7**).

Generally, the empirical impact of hypothesized health system variables and personal variables on responsiveness scores, both in individual and aggregate analysis, offers further proof of criterion validity and of the universal nature of the responsiveness concept.

Is a derived responsiveness measurement strategy feasible and valid for the Netherlands' sub-system of perinatal care?

Following the global multi-country analyses, this thesis explored three aspects of responsiveness for a sub-system of perinatal care in the Netherlands: the questionnaire's psychometric properties, the importance of domains, and the determinants of responsiveness. The Dutch perinatal responsiveness questionnaire, ReproQ, was a faithful adaptation from the pool of 28 responsiveness domain experience items, and 8 responsiveness importance items from the Multi-Country Survey Study and World Health Survey (**chapter 8**). Constructing parallel questionnaires for antenatal and postnatal care separately, the ReproQ consisted of 104 questions on responsiveness experiences (25 antenatal, 40 birth, 39 postpartum phase), 29 questions for maternal and health care characteristics and 8 importance-of-domain questions.

Overall, the ReproQ demonstrated satisfactory to excellent psychometric properties. It upheld the universality of the responsiveness concept attested to in the global datasets, as the taxonomy of responsiveness domains held its form in factor analysis (**chapters 8, 9**). Also the importance rankings were consistent with those found at the global level: for example, the "respect for persons" domains were more important than the "client orientation" domains. Communication and dignity were rated globally as most important, on average by 19% of the populations in the 41 countries (MCS), and by 26% and 22% respectively in the perinatal sub-system of the Netherlands. Prompt attention received less endorsement perhaps because waiting times were not a common problem. The stronger role for health-related characteristics in the postnatal ReproQ compared to the antenatal ReproQ can be explained from the birth consequences and confirms the adequacy of having both measurements (antenatal, postnatal). Prior to birth the great majority of women are healthy, but after delivery some women suffer from complications. The evidence of this thesis (and from later national implementation) supports the excellent translation potential of the original responsiveness concept.

CROSS-CUTTING QUESTIONS ON UNIVERSALITY

Building on the evidence from the previous section that demonstrated that the WHO responsiveness concept is as defined and can be adequately measured at the population and clinical levels, we now discuss four overarching questions.

- (i) Is it justified to have responsiveness as a separate goal (beyond health and financial outcomes), for the measurement of health systems performance?
- (ii) Is responsiveness a superior concept than hitherto used concepts and measures, in particular patient satisfaction measures?

- (iii) Given the wider concept responsiveness represents, is it possible to measure responsiveness with a reduced set of domains?
- (iv) Given the domains, is it possible to improve the core questions measuring responsiveness?

(i) Is it justified to have responsiveness as a separate goal?

Some commentators on the World Health Report 2000⁷ regarded responsiveness as a less pertinent health system goal than financial protection and health outcomes (see **chapter 1** Figure 1.1). Arguments in favour of a responsiveness goal were chiefly, originally, philosophical (the means in health care are, ethically-speaking, not secondary to the ends). Soon studies showed empirical estimates of stated preferences (individuals ranked the three main health system goals: health, responsiveness, fairness in financial contribution), which were also favourable. Respondents from 51 countries assigned larger importance weights to responsiveness than to fairness in financial protection coverage (range: 24% to 33%).⁸ This thesis provides additional evidence for responsiveness assuming a primary role in evaluating health systems, however based on the rationale that it is an important means to health. Data show that responsiveness has instrumental value for the both coverage and health outcomes; and reversely: worse health outcomes arise when there is worse responsiveness (**chapters 3, 7, 9**). While undoubtedly financial barriers have enormous repercussions for health outcomes⁹, it is together financial protection (the guarantee that essential care is always affordable) and responsiveness that predict health and health service outcomes (**chapter 7**). Knowing which aspects of the patient's experience are affected by service changes is furthermore a relevant consideration for designing and managing quality (**chapter 6**).

(ii) Is responsiveness a superior concept?

The pioneers of the many early patient satisfaction surveys deserve acknowledgement for bringing the patient perspective into the definition of health care quality.¹⁰ Their provocative struggle began a transformation of health services in particular in the UK and the USA in the 1980s. Satisfaction was a prior evaluative concept of quality introduced into health measurement, which was transferred from the field of sociology.^{11,12} These pioneers made a substantial impact on the readiness of the health services or health systems field to include user's input in the evaluation of health services, both in terms of applying evaluative concepts from the client's perspective, and using the client as superior source of information.

While patient satisfaction forms part of the ancestry of responsiveness, there is a subtle yet important difference that pertains more to the response than to the domain selection. Responses to responsiveness questions record the presence or absence of an experience: the outcome is

the experience, at the extreme it is an epidemiological incidence. The choice and phrasing of the items intrinsically reflects a universally desirable attribute. This is the meaning of expectations i.e. 'universal norms', in the original definition¹³ (a concept that is sometimes confused with individual's personal expectations).¹⁴ By contrast satisfaction responses do not state what happened, but rather elicit a degree of fulfilment of an internal standard. While initially a breakthrough reflecting patient empowerment in health system evaluation, over time the satisfaction concept has shown three major ambiguities. First, the fulfilment of an individualised standard can be psychologically interpreted as well-being. This concept is only open to interpretation as universal in terms of welfare of individuals and not in terms a depiction of services. Second, when decomposed, a satisfaction question embodies three different questions: did the stated activity done/take place? (1); what was the performance? (2); and was the outcome according to your expectation? (3). If satisfaction is high, one may assume (1),(2) and (3) are all accounted for, but if satisfaction is low, its interpretation is ambiguous, we do not know which of (1), (2) and (3) are implicated and this hampers action to improve. Third, individual expectations play a legitimate role in the satisfaction score, whereas only universal expectations (read: standards) play a legitimate role in the responsiveness score and can be adjusted if needed, as shown. In summary, the satisfaction response is an individual, non-universal, compound measure, whose intrinsic ambiguities makes it less suitable for accountable quality improvement programs. The additional options for addressing reporting behaviour bias, in our view, identify responsiveness as the choice measure, being factual and focused on universal standards for quality of services.^{15,16}

There has been a notable uptake of responsiveness-type measurements in national performance statistics and, increasingly, in clinical quality registries. In the latter context they are better known as patient reported experience measures (PREMs). In countries like United Kingdom and the United States¹⁷, patient experience metrics are part of the measurement system. In the UK, the reporting approach is oriented to the overall person experience ("Ensuring that people have a positive experience of care").^{18, 19} In the USA reports draw on surveys of the patient experience developed by AHRQ, known as "CAHPS", which WHO drew upon in the original design of its responsiveness surveys (see also: <http://www.hcahpsonline.org/home.aspx>). The implementation of responsiveness like measures under PREMs is also rapidly expanding in Europe, in particular in countries where quality registries have been set up (e.g., in the Swedish national quality registers²⁰). In view of the developmental process of quality of life measures over time, it is unsurprising that there has been abundant and organic development of local measures.

(iii) Is it possible to measure responsiveness with a reduced set of domains?

Even when the comprehensiveness of the responsiveness concept is justified by human rights and the universality of the responsiveness concept is judged as adequate by empirical study, a

relevant question remains for its practical use: is it acceptable to measure responsiveness with a smaller set of domains? This demands prioritization of domains, which is a conceptual question (whereas the related question on fewer items per domain is about measurement precision and stability).

Using the criteria of universal importance according to users' rankings, then the most important personal domains are: dignity and communication; whereas the primary setting domain is: prompt attention. A different approach is to select domains from the viewpoint of instrumental importance. Based on the first of several instrumentality argument - of being sensitive to those in poorer health (as measured by utilization of inpatient services) - a further fourth core domain could be added to existing 'settings' domains: quality of basic amenities. This domain was remarked upon most consistently, relative to other domains, by patients presenting for different types of inpatient visits (**chapter 6**). Also, despite the care taken in measurement - low average quality of facilities was more consistently remarked upon in less-developed countries (**chapter 5**), probably owing to users' perception of the need for improvement²¹, making it important for international comparisons.

Further instrumentality arguments are useful for further assessment of the reduction of domains. As regards a second instrumentality argument - impacting clinical outcomes - communication could be arguably considered a top priority, as this thesis found associations with better coverage (**chapter 3**) and other studies also testify to its instrumental importance.²²

A third instrumental argument relates to the health systems performance goal of reducing responsiveness inequality. Domains that are more sensitive to (variable across) subgroup experiences could be identified on this basis. The characteristics offered by the PROGRESS acronym²³ prove useful for a systematic checklist factors to consider for social (in)equity: place of residence (urban/rural), race/ethnicity, occupation, gender, religion, education, socioeconomic status, and social capital/resources. This thesis found evidence of equity-sensitivity for socioeconomic status (as measured by a relative ratio of permanent income (average asset index of the lowest to highest quintiles)) for the domains of dignity and communication (**chapter 6**). Absolute inequality measures (quintile 1 minus quintile 5) also supported the communication domain, but favoured the choice domain instead of dignity. Choice is a somewhat controversial alternative for other reasons as discussed below.

A fourth instrumental argument, not studied in this thesis, is the sensitivity for intentional change. So far research into the improvement of responsiveness (and its predecessor-instruments) is sparse, even at the local or clinical group level. We do not know how stable domain scores are, and how quality programs (at any level, from individual audit to international approaches like

International Consortium for Health Outcomes Measurement) exert their influence, and which domains are most susceptible over time.²⁴

Of all the domains of the WHO concept, no doubt “choice” is the most controversial, with its advocates primarily in countries with strong private health-care markets. Sensitive to the controversy, the WHO World Health Report 2000 also assigned the smallest weight to the choice domain (5%)²⁵ (page 32). As a feature of responsiveness choice has more often been hailed as instrumental variable to activate the market mechanism to produce quality services, rather than as an aspect of experience per se. Other researchers note the weaker evidence²⁶ and conceptual justification with it lacking a solid grounding in human rights (where one valid option to fulfil the right to health is enough). Also this thesis showed that more educated subpopulations favoured the choice domain, making it a regressive domain to be further disqualified on equity arguments (**chapter 5**).

(iv) Is it possible to improve the core questions measuring responsiveness?

This fourth question assumes the domains as given and goes into the scope of the question items, in particular of the four primary (core) domains. The ‘long’ face-to-face WHS used 13 to 15 items to measure the full 7 or 8 responsiveness domains in the inpatient or outpatient setting; of which 8 items to measure the 4 core domains in either setting (2 per domain).

The dignity domain (respect, physical privacy) can be enhanced by including *caring/emotional support* (e.g., “do you feel you got enough emotional support from hospital staff during your stay?” derived from the Picker-based NHS questionnaire. This aspect is a universal value in clinical care.

Communication (clear explanations, time for questions) can be enhanced by covering the aspect of *communication between doctors and, or other health professionals*; which can be referred to as a ‘coordination’ aspect (an instrumental, facility focused management issue) but also as the ‘consistency’ or ‘continuity in communication’ aspect (directly related to the client). The questionnaire items from the inpatient rheumatoid arthritis patient experiences (CQRA: Commissioning for Quality in Rheumatoid Arthritis; see: <http://www.nras.org.uk/commissioning-for-quality-in-rheumatoid-arthritis/cqra>) provide an example of addressing ‘consistency’ as follows: “are the clinic staff fully up-to-date with your health situation?”. In the ReproQ used in this thesis, the communication question refers to the aforementioned uniformity of information, which often emerges as a problem of care provision, in particular when care chains are long extending across multiple providers.

Prompt attention (wait, travel time) can be enhanced to include “convenience” beyond travel distance per se (e.g., opening hours, etc.). With the emergence of all types of internet supported services, existing questions can be adapted or new ones added. Finally, quality of basic amenities (cleanliness, space in rooms to wait or stay) can be extended to include good organization “how well organized did you find the unit” (“HowRwe” Questionnaire²⁷). This item is preferable to asking about clinical quality “how adequate/competent do you judge your doctor” as these other items no longer cover experiences, but lay interpretations of essentially specialized services.

We address only briefly the question on reduction of the number of items for the core domains, for experiences and reporting behaviour adjustment, and consider the cases of global/national versus clinical use. For global or national use, we recommend a minimum of four items: namely: being greeted and talked to respectfully (1); how clearly health care providers explained things (2); the time you waited before being attended to (3); and cleanliness of the facilities (4). For clinical use the evidence suggests more items are needed and therefore the existing items (long World Health Survey questionnaire) and the additional items suggested above should be used, in particular to identify corrective actions.²⁴

For reporting behaviour adjustments, as this thesis showed personal characteristics are more important than the specific domain responded on, it is reasonable to reduce the number of vignettes through using a single proxy vignette set that covers more than one domain. At the national level if only four domains are being measured, then one set of two vignettes describing the “good” and “moderate/bad” scale points for a single ‘personal’ domain; and another set covering the ‘setting’ domains should suffice. At the clinical level, each set of vignettes would need to be larger (at least three).

CROSS-CUTTING METHODOLOGICAL ISSUES

Two new methodological contributions were made by this thesis: multi-level analysis applied to responsiveness as a so-called PREM; and empirical assessment of heterogeneity in reporting behaviour, which rests on advanced multi-level analytical techniques.

Multi-level analysis

Multi-level analysis is essentially a determinant-outcome analysis which separates in advance relevant explanatory factors, based on a *priori* assessments of the aggregational level. The smallest level of aggregation is usually the individual response of a respondent to a question. Some determinant factors act at the group level are age, education, health status; others at the place/hospital level (which joins clients/patients visiting the same hospital; say governmental owned

or not), others at the country level (e.g., system of government, national health expenditure). Using multi-level analysis is a more careful approach to estimating the effects of determinants, where the testing of significance of an effect is mitigated by the aggregation level. For this thesis, multilevel analysis was used to estimate the determinants of reporting behaviour and the determinants of actual experiences on two levels in parallel approaches (**chapters 4, 6**), each with a set of theory-driven variables (individual-level) to decompose the group/place and country-level effects. These analyses explored a broader variable range than has previously been explored.

For determinants of reporting behaviour, the method was useful to uncover the unforeseen need for adjustment for within-country comparisons. Technically, previous hierarchical ordered probit models were limited in the number of variables they could include. As well, they did not account for clustering and the violation of the variance independence at - in our case - the country-level. Thereby, these other models understated the role of grouped (individual-level) characteristics for explaining reporting behaviour heterogeneity within countries. Our results may now serve different information users.

The determinants of experiences at the individual-level were qualitatively assessed. It is now clear that health service characteristics matter more than personal characteristics, some with larger effects than others, which can be influenced by health policy change.

The individual-level determinants for experiences are relevant to health service managers and - in the perinatal case - clinical providers, and for policy-makers concentrating on service incentives within countries. While the country-level determinants reported elsewhere²⁷ are relevant at a broader national health policy level or intercountry comparisons. (**Chapter 4**)

Measurement is expensive. In addition, the original proposed adjustment techniques for calibrating responsiveness could add to patient burden (e.g., adding vignette questions to surveys). A staged procedure, as was used in this thesis (**chapters 4, 7**), could help to reduce costs prior to comparative system performance assessment. This comment is particularly relevant at the national, policy level. Having now this thesis confirm a fair consistency of effects of individual-level characteristics on reporting behaviour (across more than 60 countries), it may be possible to anticipate the need for the inclusion of vignettes where information on the patient/user profile is already available. If profiles across institutions are similar, the heterogeneity in responses would be present but on average cancel each other out, making the collection of vignette information and the calibration process unnecessary. On the other hand, if patient profiles across comparator units are sufficiently different or unknown, collection of vignettes is needed.

Using a staged approach once vignettes are collected is also advised. In a first stage the assessment of the determinants of the reporting behaviour heterogeneity should be undertaken. In a second stage, the decision on the appropriate adjustment technique should be undertaken. It may be possible to use less computationally intensive patient-mix standardization techniques. For clinical and health service managers, responsiveness PREMs need individual-level adjustments using non-parametric or parametric techniques, which include the newer variables of education, health states and the burden of health care.

Generally, a staged approach to analyse the source of reporting behaviour bias is more conservative and recommended. As regards the source of reporting heterogeneity: ill-health as a source prompts different action as compared with education or income as the source of reporting behaviour heterogeneity.²⁸ The policy-maker or manager would have a different motivation to improve quality in service A relative to service B for the identical post-adjustment differential attributed to a different source of bias.²⁹ There would be a different imperative for action if the source of bias in the adjustment was owing to harsher negative ratings by sicker populations in service A than if owing to more confident judgements of more educated (but healthy) populations, all other things being equal.

LIMITATIONS OF THE STUDIES DESCRIBED IN THE THESIS

The key limitations across the chapters are summarized here. The WHO survey design and the perinatal studies focused on cross-sectional data collection and, for both, costs and interview burden were a consideration that resulted in limitations. Regrettably, the design limited options for testing hypotheses on (1) the effect of intentional interventions to improve responsiveness, at any level; and (2) the interaction between health and responsiveness.

Now at the time of writing (May 2018) the ReproQ has been made part of the routine quality system of two large maternity services (GJ Bonsel, personal communication, 21 May, 2018), interventions and their effects can be analyzed. This lack of evidence demonstrates that quality of care induced analysis of PREMs with the purpose of intervening is still (2018) in its infancy.

The instrumental relevance of responsiveness (or domains thereof) for health (or domains thereof) requires information on the directionality of the (observed) statistical relationship between the two concepts, be it on the national system level, or on the clinical level. In the latter case it is possible that disease/intervention quality registries with multiple measurements per patient will show more about this relationship in the coming 10 years.

A last limitation related to the multilevel technique and associated statistical computing tools, given available resources. A single estimation took easily 8 hours to run, on above averagely powered personal computers. Also the systematic interpretation of output was demanding. Still, the technique provided the desired framework for unequivocal assessment of the importance of individual-level determinants of responsiveness behaviour and experiences that were not considered in much detail previously and that had assumed most effects at the country-level.

RECOMMENDATIONS

Following a general recommendation, the specific recommendations are grouped according to four user groups: (i) international organizations; (ii) country-based policy makers and health service managers; (iii) medical professionals; and (iv) researchers.

General

Recommendation 1. Responsiveness data from service users should be routinely part of any health system evaluation, at any level. At minimum a set of 4 core domains should be covered, with, preferably, internationally standardized questions, implying a minimum 8 survey questions: 4 domain experience questions and 4 vignette questions (2 constructed for one domain of each 'personal'/'respect for persons' and 'setting'/'facilities'/'client orientation' groupings).

International

Recommendation 2. International organizations should advocate for a standard set of responsiveness PREMS to be included in measurement instruments for health and in particular in those used for assessing progress towards universal health coverage, under Goal 3 of the Sustainable Development Agenda.

Recommendation 3. Presentation of responsiveness data should correct for reporting behaviour heterogeneity including adjustments for group and country-level characteristics. Experiences should show (average) level information and distributive (equity) information. The main sources of reporting heterogeneity should be clearly labelled by country.

National policy-makers and health service managers

Recommendation 4. Responsiveness indicators need to be institutionalized in health and social services information systems. This implies measuring responsiveness across different geographic areas of the country, or other types of jurisdictions that have the mandate for health (governance) in the country (e.g., private, public). Questionnaires should be adapted to ambulatory, inpatient and combined/referral context. Increasingly useful are digital surveys that allow

for adaptive interviewing (see Adult Inpatient Survey 2015 accessed on 11 June 2016 <http://www.nhssurveys.org/survey/1619>).

Recommendation 5. Reports on responsiveness experiences should be adjusted for reporting behaviour heterogeneity. Adjustment procedures should include health state level/burden of care, education or income, sex and age. If needed, four vignettes can be used (Rec. 1).

A minimal amount of information to capture expectation differences will be needed. Data collection need not be onerous and the need for vignette information should be decided on the basis of comparing the population profiles across the units for comparison. These reports should also include information on reporting behaviour heterogeneity in order for it to be transparent (Rec. 3).

Recommendation 6. The sole use of global (total) experience scores is not advised. Each domain contains unique information, as also found elsewhere.³⁰ Aggregate scores are best used for a minimum quality threshold that is communicable to clients, providers and policy-makers alike. Aggregate analysis should be followed by assessment of particular domain outliers. Dashboards for managers by domains can help them to check their positions against national norms.

Clinical professionals

Recommendation 7. Clinical registries should be reviewed for the domains of responsiveness PREMS they cover and consideration should be given to coverage of the most important domains recommended by this thesis: dignity, communication, prompt attention and quality of basic amenities, including their recommended enhancements.

Recommendation 8. Given the suitability of contextual adaptations (described in this thesis) extension of the application into specific sub-systems defined by health specialty should be considered. Adaptations have already been used for rheumatoid arthritis, and the perinatal sub-systems (e.g., in the UK 2015 maternity survey <http://www.nhssurveys.org/survey/1559>).

Recommendation 9. Professional groups, clinicians, and managers should test, with the support of responsiveness and intervention measurement, the most effective quality cycle that improves responsiveness.

Researchers

Recommendation 10. Researchers should strive for coordination and harmonization of a generic PREM tool which potentially can be used in both population research and clinical/medical regis-

try research (like the EuroQol foundation does for the EQ5D health measure). This includes the creation of valid translations and standard analytical schemes.

Recommendation 11. Further research should be made into the association of responsiveness with service access and compliance, in particular for disadvantaged groups.

Recommendation 12. How to adapt responsiveness questionnaires to specific sub-systems, like maternal health services, rheumatology or dermatology³¹, needs further investigation. Questionnaire design needs to identify specific processes that have unequivocal interpretation as a desirable aspects of care performance. Also, researchers need to test the inclusion of items which in that sub-system may explain extreme responses. Finally the best timing of measurement and recall within the questionnaire needs specific attention for the specific sub-systems.

Recommendation 13. The feasibility of creating vignettes sets that cover more than a single domains (e.g., dignity and communication) should be investigated and their comparative efficiency and specificity in adjusting for reporting behaviour heterogeneity assessed.

REFERENCES

1. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
2. World Health Organization. State of Inequality, Geneva: World Health Organization, 2015.
3. Rice N, Robone S, Smith P. Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness. *Eur J Health Econ.* 2011;12:141–62.
4. Goddard M, Smith P. Equity of access to health care services: theory and evidence from the UK. *Soc Sci Med.* 2001;53:1149–62.
5. Shengelia B, Tandon A, Adams O, Murray CJL. Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy. *Soc Sci Med.* 2005;61(1):97–109.
6. Tanahashi T. Health service coverage and its evaluation. *Bull World Health Organ.* 1978;56(2):295–303.
7. Williams A. Science or marketing at WHO: a commentary on World Health 2000. *Health Econ.* 2001;10:93–100.
8. Gakidou E, Murray CJL, Evans DB. Quality and equity: preferences for health system outcomes. In Murray CJL, Evans DB, editors. *Health systems performance assessment.* Geneva: World Health Organization; 2003: 667–81.
9. Pearson M, Colombo F, Murakami Y, James C. Universal health coverage and health outcomes. Final Report. Paris: OECD, 2016 (22 July). Accessed on 6 April 2018 from: <https://www.oecd.org/els/health-systems/Universal-Health-Coverage-and-Health-Outcomes-OECD-G7-Health-Ministerial-2016.pdf>
10. World Health Organization. Technical consultation on concepts and methods for measuring the responsiveness of health systems. In: Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism.* Geneva: World Health Organization; 2003. p.115–24.
11. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med.* 1997;45(12):1829–43.
12. Campbell A, Converse PE, Rodgers WL. *The quality of American life; perceptions evaluations and satisfactions.* New York: Russell Sage Foundation, 1976.
13. Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bulletin of the World Health Organization* 2000;78(6):717–29
14. Mirzoev T, Kane S. What is health systems responsiveness? Review of existing knowledge and proposed conceptual framework. *BMJ Global Health* 2017;2(4):e000486.
15. Murray CJL, Kawabata K, Valentine NB. People's experience versus people's expectations. *Health Affair.* 2001;20(3):41–3.
16. Bleich S, Özaltın E, Murray CJL. How does satisfaction with the health-care system relate to patient experience? *Bull World Health Organ.* 2009;87:271–78.

Chapter 10. Health System Responsiveness by N Valentine

17. Agency for Healthcare Research and Quality. 2014 National healthcare quality and disparities report. Rockville MD: Department of health and human services, Agency for Healthcare Research and Quality; (May) 2015.
18. Department of Health: NHS. Outcomes Framework for England, 2013-2016. London: Department of Health, United Kingdom; (January, 23) 2013.
19. NHS Outcomes Framework Reports. London: Health and Social Care Information Centre; (February) 2016. Accessed on 6 April from: <https://digital.nhs.uk/catalogue/PUB30226>
20. Nilsson E, Orwelius L, Kristenson M. Patient-reported outcomes in the Swedish National Quality Registers. *J Intern Med* 2015 277(1):141–53
21. van Doorslaer E, Masseria C, Koolman X et al. for the OECD Health Equity Research Group. Inequalities in access to medical care by income in developed countries. *Can Med Assoc J*. 2006;174(2):177–83.
22. Kunrintemi V, Spatz ES, Di Capua P, et al. Patient–Provider communication and health outcomes among individuals with atherosclerotic cardiovascular disease in the United States. *Circ Cardiovasc Qual Outcomes*. 2017;10(4):e003635 (8 p.).
23. Evans T, Brown H. Road traffic crashes: operationalizing equity in the context of health sector reform. *Int J Inj Control Sa*. 2003;10(1):11–12.
24. Scheerhagen M, van Stel H, Tholhuijsen D, et al. Applicability of the ReproQ client experiences questionnaire for quality improvement in maternity care. *PeerJ* 2016;4:e2092 (16 p).
25. World Health Organization. Health systems: improving performance. The world health report 2000. Geneva: World Health Organization; 2000.
26. Costa-Font J, Zigarette V. 'The choice agenda' in European health systems: the role of 'middle class' demands. In: LEQS Paper No 82. London: London School of Economics and Political Science; (May) 2014.
27. Benson T, Potts HWW. A short generic patient experience questionnaire: howRwe development and validation. *BMC Health Serv Res*. 2014;14:499.
28. Robone S, Rice N, Smith P. Health systems' responsiveness and its characteristics: a cross-country comparative analysis. *Health Services Research* 2011:2079–100.
29. Valentine NB (2011, August). Contribution to the discussion presented at the Ordinary Meeting of the Royal Statistical Society, Dublin, Ireland. "Vignettes and health systems responsiveness in cross-country comparative analyses" by Nigel Rice and Peter Smith. Published in: *Journal of the Royal Statistical Society Series A (Statistics in Society)* 01/2012; 175(2):360–1. doi: <http://dx.doi.org/10.2307/23251886>.
30. Krol MW, de Boer D, Rademakers JJ, Delnoij DM. Overall scores as an alternative to global ratings in patient experience surveys; a comparison of four methods. *BMC Health Serv Res*. 2013;13:479.
31. van Cranenburgh OD, Krol MW, Hendriks MCP, de Rie MA, Smets EMA, de Korte J, Sprangers MAG. Consumer Quality Index Chronic Skin Disease (CQI-CSD): a new instrument to measure quality of care from the patient's perspective. *Brit J Dermatol*. 2015;173(4):1032–40.



PART IV

Additional bibliographic matter

ENGLISH SUMMARY



BACKGROUND AND INTRODUCTION

This thesis discusses the theoretical and empirical aspects of the World Health Organization (WHO)'s "health system responsiveness" (hereafter: responsiveness) concept. The responsiveness concept formed part of the set of three WHO health system goals proposed in the 2000 Health Systems Performance Assessment framework, which also included health outcomes and fairness in financing goals. These universal health system goals were developed as part of a global programme focused on using universal and comparable metrics of health and other health system related variables.

The responsiveness of a health system as a concept was defined by WHO and leading scientists as the health system's ability to meet the universal, legitimate expectations of its users (or clients, patients) with regard to non-medical aspects of the way they are treated and the environment (or setting) within which they are treated. Responsiveness as measured replaced a wide variety of pre-existing measures of process and organisational features - the heritage of Donabedian's prevailing quality approach - by a unified system of asking patients and clients for their experiences, as a single yet sufficient source of information.

Thus the WHO measurement logic for responsiveness follows the standard logic for measuring self-reported health outcomes (or quality of life) through structured questionnaires, with a set of well-chosen health domains (e.g., pain, mobility), and associated questions with suitable response options. For measuring responsiveness, questions ask respondents to rate their most recent experience or contact with the health system, where 'experience' is described by multiple experience questions (or 'items') specific to each of 8 domains. This domain-with-associated-items approach is similar to the technique used in the health status instruments like the SF-36, where multiple items are used to cover a single domain. For responsiveness, the questionnaire items are also contextualized for inpatient (overnight stay) versus ambulatory or outpatient experiences in different sections of the questionnaire. Depending on the last experience, the interviewee responds to either the inpatient or outpatient sections of the survey.

Of the 8 responsiveness domains, 4 were grouped as 'personal' or "respect for person(s)" domains: autonomy, communication, confidentiality, and dignity ; and 4 were grouped as 'setting' or "client orientation" domains: choice, prompt attention, quality of basic amenities and (access to) social support. The WHO undertook the global collection of responsiveness data, together with other data, between 2000- 2004 through two rounds of multi-country household surveys called the Multi-Country Survey (MCS) Study (2000-01) and the subsequent World Health Survey (WHS) (2002-03/4). The survey questionnaires were constructed of thematic modules. Apart from responsiveness, the survey addressed health status, features of the health system

including coverage and background information such as health insurance status, health expenditures, socio-demographics and income. For the broader goal of synthesizing information into overall comparative country assessments, it also included so-called health state valuations, as well as respondent views on health system goals.

The MCS and WHS shared many similarities but there were some notable changes in the WHS: extending the number of modules with more face-to-face, long (90 minute) interviews; using more surveys with sample weights; including the experiences of children up to twelve (reported through a parent as recommended by the WHO expert consultation on measuring responsiveness); reducing the number of items on each responsiveness domain; and expanding the items on health service characteristics.

This thesis draws upon data from interviewer-supported responsiveness questionnaires made publicly available by WHO (106 surveys from MCS and WHS). The earlier thesis papers (chapters 2, 5) use the MCS data that resulted from the first wave of surveys. After a comparison of the two questionnaires (chapter 3), the latter papers (chapters 4,6,7) take advantage of the wider WHS set of health systems items. In addition, a separate study, called the Responsiveness in Perinatal and Obstetric Health Care Questionnaire (ReproQ) study adapted the WHO responsiveness questionnaire to perinatal care in the Netherlands. The first phase was undertaken between 2009-11 and two papers are based on this study (chapters 8,9). Including the Netherlands study, this thesis analyses about 260,000 client interviews on responsiveness from 107 surveys covering 83 countries of all levels of development.

THESIS OUTLINE

Following [Chapter 1](#), the introduction, the thesis papers are grouped into three parts. The first part (chapters 2-4) examines the psychometric properties i.e. feasibility, reliability and validity of the responsiveness questionnaires used. Chapter 2 appraises the MCS responsiveness experience items with a psychometric analysis; the second paper appraises the responsiveness concept of 'access to acceptable care'. Chapter 2 also compares the psychometric properties of the improved WHS questionnaire, comparing it with the MCS. Using a multi-level analysis, the third paper reviews the intricate issue of reporting behaviour variation or bias (or reporting behaviour heterogeneity), the phenomenon that respondents using response scales with the same words (e.g., "very good" or "very bad"), or scales consisting of a calibrated line ("thermometer") to mark with a cross (so-called visual analogue scales), do not have the same level of achievement or severity in mind when responding.

The second group of papers (chapters 5-7) addresses what universal factors determine the responsiveness of any health system. The first paper reviews the influence of personal and setting characteristics on the individually ranked importance of the 8 responsiveness domains; the second paper employs multi-level analysis of responsiveness determinants grouped according to policy relevance (personal, service, and national/country)). The third paper uses data aggregated at the country level (ecological analysis) to test the association between the attainment of health system responsiveness and the attainment of other accepted health system goals. Here, responsiveness is viewed as a means to better health and health systems functioning.

The third group of papers (chapters 8-9) shows the psychometric evidence of the adapted responsiveness questionnaire and the determinants of responsiveness in the context of perinatal care in the Netherlands.

The final part of this summary presents the thesis discussion and recommendations.

Responsiveness as a measurable and universal concept: chapters 2-4

Chapter 2 evaluates the psychometric properties of questions on health systems responsiveness using 41 interviewer-administered MCS surveys. It evaluated the feasibility, reliability, and construct validity of measuring responsiveness domains using 33 items with ordinal polytomous response options. To address universality it compared responses from populations identified by countries, sex, age, education, health and income. Missing rates were acceptable. Intertemporal reliability was acceptable in 6 (of 10) sites conducting retests, where Kappas (K) ranged from 0.54 to 0.79, but low in 4 sites ($K < 0.5$). K was higher for male, educated and healthier subpopulations. The 8-domain structure of the responsiveness concept was confirmed by factor analysis. Criterion validity ('known group comparisons') was evidenced by higher income of individuals being associated with more positive responsiveness reports. The paper concluded that quality of care issues addressed by WHO's responsiveness questions are understood and reported adequately across diverse populations (universality).

Chapter 3 elaborates on the theoretical aspects of the responsiveness concept and associated measurement strategy and presents a psychometric analysis of responsiveness in the WHS, making comparisons with the MCS as the WHS aimed at some improvements. Arguments are put forward for placing responsiveness in the portfolio of 'equity in access' concepts. In both the MCS and WHS, psychometrics were favourable. Despite increased length, for feasibility, the average item missing rates did not exceed 3.25% in the WHS, nearly 1% lower than in the MCS. The distribution of responses over the 5-point response scale was improved; temporal reliability averaged for 53 countries showed moderate reliability (0.4-0.6). Comparing like-with-like kappa statistics from the MCS and WHS was only possible for two countries (India, China) and this

comparison showed improved kappa statistics for both countries (for the WHS: 0.7). The overall validity of the concept was supported by the exploratory factor analysis in both the MCS and WHS. Internal validity metrics were acceptable and a little better for the (longer) WHS. Criterion validity ('known group comparison') was confirmed by the expected social gradient in responsiveness scores between and within countries. Also, average country level responsiveness correlated positively with antenatal care coverage as anticipated. The paper concluded that measuring responsiveness in both the MCS and WHS was valid, though a little better in the WHS.

Chapter 4 addresses the literature on reporting behaviour bias, or reporting behaviour 'heterogeneity' (also called 'response behaviour') which refers to the differential use of a response scale. Reporting behaviour bias can be measured and the data subsequently corrected if a survey contains extra calibration questions included for the purpose of this correction. The usual approach is to create 'vignettes', i.e. formal descriptions of hypothetical situations or scenarios. A vignette is a short paragraph in the questionnaire that is read out to the respondent (as stimulus). It depicts a 'scenario' of a patient in a health care setting experiencing a certain level of care (e.g., waiting for admission, receiving explanations on diagnosis and treatment), typically recounted as a third person's experience. It is pre-designated to depict all experience levels from very bad to very good experiences on a particular issue, without disclosing to the respondent the envisaged level, and thus allowing the respondent to scale the experience depicted. By using multiple vignettes depicting scenarios across the scale, and comparing responses across respondents, one may derive for each respondent, or group of respondents, the use of the scale relative to the average of all respondents. In other words: the numbers given to a set of hypothetical scenarios by each respondent enable to interpret the number this respondent gives to his or her own unshared experience. Thus, one may recalibrate all responses to a common scale, which in turn allows for considerably less biased comparisons within and between countries.

This chapter used the WHS dataset of about 150,000 respondents in 64 countries, which contained 5 calibration vignettes for each of 8 domains. Using a multi-level ordered probit regression analysis, chapter 4 describes response patterns according to country and individual features. The multilevel model assumed two levels, individual and country, and tested response heterogeneity arising from the country level (e.g. social norms) and from the individual-level (e.g., individual educational attainment). Country-level patterns are described by the intraclass correlation coefficient by vignette and domain, while individual-level patterns are described in terms of a) contraction/elongation and b) shift (up, down) of particular parts of the scale. Country effects were found to be more important for explaining heterogeneity in responses for (ordered by effect size) quality of basic amenities, prompt attention, and confidentiality domains.

The patterns attributed to individual-level variables were that scale elongation/contraction (a) occurred with the variables sex, education, caring for others with chronic illness and the importance of responsiveness (values). A shift effect (b) was most strongly associated with own health, with poor own health giving rise to lower scores on scales used (implying higher own responsiveness scores after correction). A critical finding was that patterns were fairly constant across the 8 responsiveness domains. This study showed that reporting behaviour is related to individual characteristics much more strongly and consistently than thought before, with still sizeable country effects. These features not being evenly distributed, it confirmed the need for the described correction approach, accounting for country and individual-level sources of bias in comparative reporting on responsiveness (and other outcomes), and in particular for inequity research. It was reassuring that across countries, vignette responses were sensitive to a small set of hypothesis-based variables that were measurable and therefore could be used within and between country adjustments.

Universal determinants of responsiveness : chapters 5-7

Chapter 5 investigates whether people assign different importance to the 8 responsiveness domains, where the starting point of the questionnaire was 8 equally important domains, throughout the globe. For this purpose, the MCS contained ranking questions on responsiveness domains. Data were used from 105,806 respondents to the MCS in 41 countries. Multinomial logit regression models were used for the questions on the most and least important domains. The variations in domain importance were explored by country level variables (country of residence, human development, health system expenditure, and 'geographic zones') and by individual-level variables (sex, age, education, (own) health status, and utilization (inpatient, outpatient)). Most frequently, respondents selected prompt attention as the most important domain. Dignity was selected second, followed by communication. Access to social support networks was identified as the lowest ranked domain. There was overall consistency between separate questions in the most and least important domain, with the exceptional reversal of the importance of dignity and prompt attention. In general, ranking convergence was stronger within countries than across countries. The strongest individual-level predictors of rankings were health status and utilization of health services, with those in better health or only having ambulatory (versus inpatient experiences) more likely to select dignity, choice and autonomy as more important relative to prompt attention. If you are ill, the thing that matters is to be served. The country-level deviations in priorities for responsiveness were most strongly expressed for countries with lower development indices, which assigned slightly higher importance to the quality of basic amenities domain; and for non-European countries, which assigned more importance to access to social support. Choice was, remarkably, relatively more favoured in Central and Eastern Europe (CEE)/ former Soviet Union (FSU) countries, and confidentiality in Eastern Mediterranean and EEC/FSU countries. Yet even with these observations, an overall ranking prevailed represented by the prompt

attention, dignity and communication domains. With these findings taken together, the ranking does not presuppose the irrelevance of the other domains that were not ranked first, considering that the questionnaire only asked for the most and least important domains. The findings do however suggest how to reduce the questions if needed, how to set priorities during health reform processes or benchmarking, and how to prioritize n cross-country comparisons. Perhaps the most important result of the many underlying data analyses was the scientific support for a universal interpretation of the responsiveness concept.

Chapter 6 asks if there are features of health systems and personnel that universally produce better or poorer responsiveness. Politically, this may be interpreted as a sensitive question, especially once we assume - see previous chapters – that this comparison is not overwhelmed by reporting behaviour biases. The WHS dataset enabled the use of a broad set of health service determinants (12 variables), measured at the individual level, and personal determinants (8 variables); multilevel regressions were developed, for each domain and inpatient/outpatient combination separately. The WHS dataset used for analysis included 120,000 respondents from 49 countries. A responsiveness frequency score was derived for each domain after dichotomizing the original 5 ordinal response categories for the each individual (poor responsiveness = "moderate", "bad" or "very bad"). This enables the computation of frequencies on a group level, and the use of exploratory regression techniques. The most striking result was the presence of a set of health service characteristics explaining responsiveness all over the globe, which were much more explanatory than personal characteristics. The following service characteristics mattered most and worsened responsiveness across multiple domains: discriminatory attitude of health workers; inaccessability in terms of measured travel times; and the service being managed by the government (rather than non-governmental bodies). For personal characteristics, rather than age and sex (the obvious literature candidates), poverty and health status were found to be important, confirming earlier descriptive findings of responsiveness inequalities (chapter 3). These (and other) observations show responsiveness to be an equity-sensitive outcome. Additional results suggested that specific pathways operated in particular domains studied.

Chapter 7 presents ecological regression analyses at the country level, using determinant and outcome variables at that level (e.g., national coverage rates for vaccination, child mortality rates). The adopted approach assumed responsiveness to be instrumental (intermediary or mediating variable) between access and health outcomes (including coverage and health status 'outcomes'). Several different regression models (negative binomial, log-linear and ordinary least squares) were compared to find the best model fit for these types of variables. Health status and coverage variables included those for noncommunicable diseases, communicable diseases, and for reproductive and child health. The results showed that dignity and prompt attention were associated with maternal mortality and TB mortality; dignity was significant in

models predicting aggregate child mortality. Here the causal path is perhaps bidirectional, rather than the more common unidirectionality. For an example of unidirectionality, a striking example of measles vaccination inequality seems to be explained by both financial protection and responsiveness, indicating the relevance of 'acceptability' and 'affordability' access for securing more intervention equality between the wealthy and the poor. These key associations found at the country level converge with similar individual-level associations (e.g., by income), reducing the likelihood of an ecological fallacy. Additionally, the relevance of both responsiveness and financial protection offers further evidence in support of the validity of the 2000 WHO framework for "Health System Performance Assessment".

Responsiveness measurement adaptation to perinatal care: chapters 8-9

The last two papers describe an adaptation of the responsiveness concept and the questions used in the global surveys to the specific context of perinatal care. The Dutch perinatal responsiveness questionnaire as described here, was created in a collaborative effort with an investigator's team of ErasmusMC, Rotterdam and University Medical Centre Utrecht. The interviewer-supported ReproQ as described here adapted WHS and MCS responsiveness items for three phases of perinatal care: antenatal (25 items), birth (40 items), postpartum (39 items) phases; asked about the importance of domains; and recorded variables on health services characteristics (e.g., referral as example of discontinuity of care) and the experience (outcomes) of the interaction.

Chapter 8 assesses the psychometric properties of feasibility, construct validity, and discriminative validity in a sample of Dutch women surveyed (n=171 respondents) through face-to-face interviews conducted 2 weeks post partum (interviews lasting between 20 and 40 minutes). The overall item missing rate was 8%. Mean Cronbach's alphas for the antenatal, birth and postpartum phase were: 0.73, 0.84, and 0.87 respectively. Within the antenatal care, birth care and post partum phases, the 8 factors explained 69%, 69%, and 76% of variance respectively for the the antenatal, birth and postpartum phases, a good result. Overall responsiveness scores were higher (better) for women whose children were not admitted. This confirmed discriminative validity and the hypothesis that dissatisfaction with health outcomes is transferred to the patient's judgement on responsiveness of the perinatal services. Overall, the ReproQ interview-based questionnaire demonstrated satisfactory psychometric properties to to discriminate between different quality levels of perinatal of care.

Chapter 9 produced an extensive set of results on responsiveness from the ReproQ survey. It focused on comparing the facility-based antenatal and delivery phase of care. Responsiveness experiences, were, as before (chapter 6), reduced to a dichotomous variable (poor responsiveness

= “moderate”, “bad”, “very bad”). The domain importance rankings were consistent with those found at the global level (chapter 5). The frequency of poor responsiveness experiences ranged from 5.9% to 31.7% for the antenatal phase and from 9.7% to 27.1% for the delivery phase. Overall for both phases, “respect for persons” (e.g., dignity, communication) domains performed better and were judged to be more important than “client orientation”/‘setting’ domains (e.g., prompt attention, quality of basic amenities). A set of expected independent variables on health services and users’ personal background characteristics explained responsiveness. On the whole, responsiveness was explained more by health-care and health-related issues than by personal characteristics as found previously (chapter 6). ‘Known group comparisons’ suggested validity: as hypothesized, poor obstetric history and an adverse birth event clearly negatively influenced responsiveness results. The evidence supported the excellent adaptation potential to specific clinical settings of the original responsiveness concept and its measurement strategy.

CROSS-CUTTING DISCUSSION THEMES

After its first presentation, some commentators on the WHO World Health Report 2000 questioned the relevance of responsiveness. At that stage, psychometric evidence like that described in this thesis was minimal. Our results justify responsiveness as a separate indispensable goal (beyond health and financial outcomes) for the quantitative assessment of health system performance. We regard as convincing the evidence on the importance of both the universality of responsiveness and on measuring the quality of the care process at an individual level. Its hypothesized instrumental relevance for health outcomes, including for universal health coverage and health equity was demonstrated.

Responsiveness entered into health system performance discourse in 1999-2000, and owes much to the older concept ‘patient satisfaction’. Yet there are subtle but important differences between responsiveness and patient satisfaction that pertain to the response rather than to the domain selection and to the derived items. Replies to responsiveness questions record the actual presence or absence of an experience while satisfaction records an internal state or emotion. Dissatisfaction may therefore refer to multiple conditions; apart from the intended quality of the service, it may point to lack of the experience when it was expected or needed, or to unmet individual unmeasured expectations. This plurality of interpretation is a major drawback for using patient satisfaction in quality assessment processes. With responsiveness, as elaborated here, the interpretation of response is more straightforward, and expectations play in view of the factual frequency-like question a smaller role, which can be corrected if needed. The increased uptake of responsiveness-type measures, labelled ‘patient reported experience measures’ (PREMs), including in clinical registries, and across medical specialties, is further testimony to the universal relevance of the concept.

Measuring responsiveness through client surveys appears robust to different settings and constraints. Certain available strategies can reduce measurement costs, while others may enhance survey sensitivity. The thesis shows that it may still be possible to obtain useful measurement results when the responsiveness domains are reduced from 8 to 4, core, domains, namely: prompt attention, dignity, communication and quality of basic amenities.

Regardless of the number of domains, vignettes should be included in measurement strategies to correct for response heterogeneity. Fewer than we did could possibly be used, given similar patterns observed across domains (on the individual level in particular). Measurement sensitivity could be enhanced by adding the following items to the 4 core domains: convenience in prompt attention; caring attitudes in dignity; communication between health professionals in communication; and extent of organization in quality of basic amenities.

RECOMMENDATIONS

The general recommendation is for any health service to record, on a routine base, individual responsiveness data. This assumes the permanent relevance and utility of quality evaluations. International organizations should advocate for a standard set of responsiveness items in measurement instruments in particular in those already being used for assessing progress towards universal health coverage, under Goal 3 of the Sustainable Development Agenda which guides the work for WHO, as do other goals for other United Nations Agencies. Responsiveness data presentations should correct for reporting behaviour heterogeneity, unless data show this step can be skipped. Within countries, clinical quality registries should consider the application of vignette-based correction.

In systems for quality reporting, aggregate global responsiveness ratings are best used for minimum quality thresholds and national benchmarking. The practice of solely reporting global scores is, however, not advised as each domain contains unique information, containing partially specific causal pathway information that is relevant to designing improvements. Adjusting for differences in reporting behaviour bias will often be needed in this context, usually requiring explicit attention and modest additional resources.

Given the ease of contextual adaptations of the responsiveness questionnaire as described in this thesis, additional extensions to specific sub-systems, defined by health speciality, are advocated. Adaptations have already been used for rheumatoid arthritis and the perinatal sub-systems (e.g., in the UK 2015 maternity survey). Professional groups, clinicians, and managers should test, with the support of responsiveness and intervention process measurement, the most effective quality cycle that improves responsiveness. Researchers should strive for the co-

ordination and harmonization of generic questionnaires (PREMs, PROMs), and conduct further research into the reduction of vignettes; finally, the association of responsiveness with service access and compliance should be investigated, in particular for disadvantaged groups.

DUTCH SUMMARY

NEDERLANDSE SAMENVATTING



INLEIDING

Dit proefschrift gaat over de theoretische onderbouwing van en empirische ervaring met het zgn. "health system responsiveness" (hierna: responsiveness¹) concept, dat door de Wereldgezondheidsorganisatie (World Health Organization, WHO) is ontwikkeld. Het responsiveness concept is 1 van de 3 doelen waarnaar een gezondheidszorgsysteem moet streven, volgens de WHO. Dit evaluatiemodel werd in het jaar 2000 geformuleerd en staat bekend als het "Health Systems Performance Assessment framework" van de WHO. Naast responsiveness, vormen gezondheid en financiële rechtvaardigheid (betaalbaarheid) de andere 2 doelen. Deze universele doelen van een gezondheidszorgsysteem werden ontwikkeld in een wereldwijd WHO project, dat beoogde om universele, tussen landen vergelijkbare gezondheids(zorg)statistieken te maken.

Responsiveness van een gezondheidszorgsysteem, als idee, werd gedefinieerd als het vermogen om tegemoet te komen aan algemeen aanvaarde, gerechtvaardigde verwachtingen van een gebruiker (burger, cliënt, patiënt) waar het gaat om de niet-medische aspecten van behandeling en dienstverlening, daarbij inbegrepen aspecten van de setting van zorgverlening. Responsiveness, als een hiervan afgeleide maat, verving een groot aantal bestaande, heterogene, en niet op elkaar herleidbare maten, die te maken hadden met kenmerken van het zorgproces en de zorgorganisatie - een erfenis van de tot dan toe dominante kwaliteitsbenadering van Donabedian. In plaats daarvan bood responsiveness een eenduidig meetsysteem, waarin aan patiënten en cliënten naar hun concrete ervaring werd gevraagd, wat als enige en toereikende bron van informatie over de niet-medische kwaliteit van zorg werd beschouwd.

De WHO benadering van het meten van responsiveness is geheel conform de standaardbenadering van het meten van zelf-gerapporteerde gezondheid c.q. kwaliteit van leven, namelijk via vragenlijsten. Deze gestructureerde vragenlijsten berusten op een onderbouwde keuze van gezondheidsdomeinen (zoals bv. pijn, mobiliteit), met daarbij gekozen vragen ('items') en bijpassende antwoorden ('responsvormen'). Bij het meten van responsiveness wordt aan de respondent gevraagd stil te staan bij de meest recente eigen ervaring met het gezondheidszorgsysteem, en deze te scoren; bij elk responsiveness domein (er zijn er 8) is een aantal bijpassende ervaringen geformuleerd, die bij ieder contact met de zorg respect voor privacy, en bejegening; relevant zijn (bv. bejegening of wachttijd). Deze domein-met-bijpassende-vragen

¹ Het begrip 'responsiveness' kent geen Nederlands equivalent. De auteur van het begrip 'responsiveness' koos met opzet voor dit begrip, in plaats van 'kwaliteit van zorg' of 'proceskwaliteit', om enerzijds het contrast met het begrippenkader van Donabedian duidelijk te maken, en anderzijds te benadrukken dat het gaat om wat de individuele client merkt van goede proceskwaliteit c.q. het voldoen van de dienstverlening aan gerechtvaardigde verwachtingen (persoonlijke mededeling Murray). Vanwege deze specifieke betekenisverlening is het begrip onvertaald gelaten.

techniek is gelijk aan die welke wordt toegepast bij gangbare gezondheidsvragenlijsten zoals de SF-36 vragenlijst: ook daar worden meerdere vragen gesteld, die bij elkaar een enkel - in dat geval - gezondheidsdomein omvatten. Voor de responsiveness vragenlijst moeten vragen worden aangepast aan de zorgsituatie, in het bijzonder of er sprake was van zorg en verblijf in een ziekenhuis, dan wel sprake van behandeling en zorg op een polikliniek of praktijk, zonder opname. De respondent beantwoordt afhankelijk van de laatste ervaring de set klinische vragen of de set ambulante vragen.

Vier van de 8 responsiveness domeinen worden gerekend tot 'persoonlijke' domeinen: (respect voor) autonomie, communicatie, respect voor privacy, en bejegening; de andere 4 zgn. 'setting' domeinen betreffen zorgtoegankelijkheid, kwaliteit voorzieningen, keuzevrijheid, en (toegang tot) sociale steun².

De WHO was in de periode 2000-2004 verantwoordelijk voor de mondiale verzameling van individuele responsiveness gegevens samen met andere gegevens. Deze dataverzameling kende 2 survey-rondes waarbij in een groot aantal huishoudens in tal van landen een modulair samengestelde vragenlijst³ werd afgenomen; de eerste was de Multi-Country Survey (MCS) (2000-01), de volgende de World Health Survey (WHS) (2002-03/4). Naast een module voor responsiveness, bevatten de vragenlijsten modules over de (eigen) gezondheid, het gezondheidszorgsysteem, ziektekostenverzekering, gezondheidsuitgaven en eigen bijdragen, en persoonlijke achtergrondinformatie, waaronder het huishoud-inkomen. Voor het specifieke doel om oordelen en scores tussen landen vergelijkbaar te krijgen, waren zgn. gezondheidswaarderingsvragen opgenomen, en vragen over het relatieve belang van verschillende doelen van een gezondheidszorgsysteem.

De vragenlijsten van de MCS en de WHS waren in veel opzichten vergelijkbaar, maar in de WHS waren enkele veranderingen aangebracht. Het aantal modules was groter, het aandeel lange face-to-face interviews (90 minuten) was groter; relatief meer surveys werden uitgezet met gebruik van steekproef gewichten bij de berekening van totaalscores; de ervaring van kinderen tot 12 jaar werd nu ook meegenomen via de ouder als proxy; het aantal items per res-

2 De oorspronkelijke domein-labels zijn: *autonomy, communication, confidentiality, dignity; en prompt attention, choice, quality of basic amenities, (access to) social support.* Wij kozen de meest gebruikte domeinnaam die in Nederland bij de betreffende items wordt toegepast.

3 Het begrip 'survey' wordt onvertaald gelaten, een term waarvoor geen exact Nederlands equivalent bestaat. Survey verwijst naar de inhoud van een vragenlijst (de items, de responsvormen), naar het proces van de vragenlijst afnemen, en soms naar het gehele gegevensbestand dat met de vragenlijst is verworven. In het laatste geval verwijst het hier naar de antwoorden afkomstig van 1 land.

ponsiveness domein was minder; en er werden meer vragen gesteld over kenmerken van het gezondheidszorgsysteem.

Dit proefschrift gebruikt de gegevens uit de interviewer-ondersteunde vragenlijsten en bijbehorende gegevens, zoals die door de WHO zijn verzameld en openbaargemaakt (106 surveys afkomstig uit de MCS en de WHS). Hoofdstuk 2 en 5 gebruiken de MCS surveys uit de eerste ronde. Na de vergelijking van de vragenlijsten van MCS en WHS (hoofdstuk 3), worden in de hoofdstukken 4, 6, en 7 de WHS vragenlijsten gebruikt, die over meer gegevens over het gezondheidszorgsysteem beschikken. In een afzonderlijke studie (de ReproQ studie) werd een responsiveness vragenlijst voor de geboortezorg ontwikkeld en gebruikt. De eerste fase daarvan vond plaats tussen 2009-11, waarin de WHO responsiveness survey aangepast werd aan de geboortezorg in Nederland⁴. Twee artikelen zijn hierop gebaseerd (hoofdstukken 8 en 9). De Nederlands studie inbegrepen, gaat deze studie over in totaal 260.000 individuele antwoorden van cliënten, afkomstig van 107 surveys, verzameld in 83 landen met allerlei niveaus van economische ontwikkeling.

LEESWIJZER BIJ PROEFSCHRIFT EN DEZE SAMENVATTING

Na de inleiding ([hoofdstuk 1](#)), volgen de overige hoofdstukken, die in 3 groepen uiteenvallen. Hoofdstukken 2-4 onderzoeken de psychometrische eigenschappen van de responsiveness vragenlijst, onderscheiden in haalbaarheid, betrouwbaarheid en validiteit. Hoofdstuk 2 onderwerpt responsiveness items zoals opgenomen in de MCS aan een psychometrische analyse. Hoofdstuk 3 gaat dieper in op de relatie van het responsiveness concept met het bredere idee van 'gelijke toegang (voor iedereen) tot minimaal voldoende zorg'. Hoofdstuk 3 vergelijkt de psychometrische eigenschappen van de MCS vragenlijst met die van de WHS, waarbij in de laatste enige wijzigingen waren aangebracht. Hoofdstuk 4 onderzoekt met multilevel analyse de ingewikkelde kwestie rond 'respons heterogeniteit'. Dit betekent zoveel dat respondenten responschalen met ankerwoorden (b.v. "heel goed" of "heel slecht") of bestaande uit een lijn met streepjes ("thermometer") waarop een kruisje moet worden gezet, verschillend gebruiken (zgn. visueel analoge schalen). Zij hebben niet hetzelfde niveau van (dys)functie of ernst in hun hoofd wanneer zij met hetzelfde woord op dezelfde plaats op de lijn hun antwoord aangeven.

De tweede groep (hoofdstukken 5, 6, en 7) onderzoekt of, en zo ja welke, universele factoren de responsiveness van een willekeurig gezondheidszorgsysteem bepalen. Hoofdstuk 5 toont

⁴ De in dit proefschrift opgenomen interviewer-ondersteunde ReproQ vragenlijst, was de pilot versie van de later ontwikkelde ReproQ vragenlijst voor zelfbeantwoording zoals deze in 2014 definitief is vastgesteld. Het responsiveness concept en de domeinstructuur zijn daarbij ongewijzigd gebleven. Verdere bespreking van deze doorontwikkeling valt buiten dit proefschrift.

welke persoonlijke en omgevingskenmerken van invloed zijn op de door een individueel toegekende rangorde van responsivens domeinen. Hoofdstuk 6 gebruikt multilevel analyse om de rol van determinanten van (waargenomen) responsivens vast te stellen, waarbij de determinanten gegroepeerd zijn naar politieke relevantie (persoon-gerelateerde, gezondheidszorgsysteem gerelateerde, en landgebonden determinanten). Hoofdstuk 7 analyseert responsivens gegevens, geaggregeerd op land-niveau (ecologische analyse) om te toetsen of er samenhang is tussen de hoogte van de gemeten responsivens, en de prestaties wat betreft andere doelen. In deze analyse wordt verondersteld dat responsivens een middel is richting betere gezondheid en een betere andere prestaties van het systeem.

De derde groep (hoofdstukken 8-9) toont de psychometrische eigenschappen van de responsivens vragenlijst die voor de Nederlandse geboortezorg werd afgeleid, en beschrijft verder de bepalende factoren van responsivens in de context van Nederlandse geboortezorg.

Het laatste hoofdstuk 10 presenteert de discussie en aanbevelingen.

Responsivens als universeel en operationaliseerbaar begrip: hoofdstukken 2-4

Hoofdstuk 2 evaluaert de psychometrische eigenschappen van de responsivens items zoals die in de 41 interviewer-ondersteunde surveys van MCS zijn gebruikt. De haalbaarheid (feasibility), betrouwbaarheid, en validiteit wordt bepaald van de 33 items met ordinale meerkeuze antwoorden voor de 8 domeinen. Universaliteit werd onderzocht door antwoorden te vergelijken naar land, sexe, leeftijd, opleidingsniveau, gezondheidsniveau en inkomen. Het aantal missende waarden was aanvaardbaar laag. Test-hertest betrouwbaarheid was voldoende in 6 (van 10) plaatsen waar dit werd bepaald; de kappa (K) varieerde daar van 0.54 tot 0.79, op 4 plaatsen lag K onder 0.50. K was hoger (beter) voor mannen, hoger opgeleiden, en gezondere subpopulaties. De 8-domein structuur van het responsivens concept werd bevestigd via factor analyse. Criterium validiteit ('known group comparisons') werd ondersteund door het feit dat hoger inkomen samenhang met hogere gerapporteerde responsivens. Het hoofdstuk concludeert tenslotte dat kwaliteit van zorg aspecten waaraan gerefereerd wordt in de WHO's responsivens items, goed worden begrepen en beantwoord over de verschillende subgroepen van respondenten (universaliteit).

Hoofdstuk 3 werkt de theoretische aspecten uit van het responsivens concept en van de gekozen operationaliseringsstrategie. Het bevat een psychometrische analyse van de (verbeterde) responsivens vragenlijst van de WHS in vergelijking met die van de MCS. Argumenten worden naar voren gebracht om responsivens te beschouwen als onderdeel van 'gelijke toegang tot zorg' concepten. In zowel de MCS als de WHS waren de psychometrische kenmerken

gunstig. Ondanks de toegenomen lengte was het gemiddeld percentage missende waarden (haalbaarheid) in de WHS niet hoger dan 3.3%, wat bijna 1% minder is dan dat in de MCS. De antwoord-verdeling over de ordinale 5-puntsschaal was verbeterd in de WHS. De gemiddelde test-hertest betrouwbaarheid van de WHS kon in 53 landen worden bepaald en was redelijk (0.4-0.6). Alleen in India en China kon een zuivere vergelijking van de K van de MCS en de WHS plaatsvinden en deze vergelijking toonde een verbeterde K aan in beide landen (voor de WHS: 0.7). De globale validiteit van het responsiveness concept werd ondersteund door 'exploratory factor analysis' van zowel de MCS als WHS responsiveness gegevens. Interne validiteit was voldoende, en enigszins beter in de (langere) WHS. Criterium validiteit ('known group comparisons') werd ondersteund door de bevestiging van de verwachte sociale gradiënt in responsiveness scores tussen en binnen landen. Ook hing responsiveness op land-niveau positief samen met antenatale zorgdekking, zoals verondersteld. Het hoofdstuk concludeert dat responsiveness zowel in de MCS als de WHS valide werd gemeten, in de WHS nog iets beter.

Hoofdstuk 4 behandelt de literatuur over antwoordstijlen⁵, wat betrekking heeft op individuele verschillen tussen respondenten in het gebruik van een antwoordschaal. Iemands antwoordstijl kan empirisch worden vastgesteld, waardoor vervolgens de antwoorden op die schaal kunnen worden gecorrigeerd. Voorwaarde is dat de vragenlijst extra calibratie items bevat, die dus uitsluitend ten behoeve van deze correctie zijn opgenomen. De gebruikelijke aanpak is om zgn. 'vignetten' op te stellen, dat zijn geformaliseerde beschrijvingen van hypothetische situaties (scenarios). Een vignet is een korte tekst in de vragenlijst die - hier- als stimulus wordt voorgelezen aan de respondent. Het beschrijft de ervaringen van een hypothetische patiënt in een concrete gezondheidszorgsituatie via een soort verhaaltje ('scenario'), waarbij de onderzoeker van tevoren een bepaald niveau van zorg wat betreft 1 of meer van de responsiveness domeinen heeft aangebracht (b.v. het moeten wachten voor opname in het ziekenhuis, het vragen en krijgen van uitleg over diagnose en behandeling). Het beoordelingsperspectief is dat van de derde persoon. De scenarios zijn zo ontworpen dat impliciet alle niveaus van de verschillende responsiveness schalen erin voorkomen, maar zonder evaluerende woorden die het niveau benoemen, waardoor de respondent uiteindelijk een score moet toekennen op de aangeboden schaal. Door verschillende gevarieerde vignetten te gebruiken, kan voor iedere respondent, of groep van respondenten, het gebruik van de schaal worden bepaald relatief ten opzicht van het gemiddelde van allen. Door van hypothetische situaties - die iedereen moet beoordelen - te bepalen welk cijfer daarvoor wordt gegeven, kan dus een gegeven cijfer aan de eigen unieke situatie in

⁵ Wij kozen voor de vertaling 'antwoordstijl' als beste benadering voor het begrip 'respons heterogeniteit' waarvoor in het Engels naast 'reporting heterogeneity' ook soms 'reporting behaviour' of daarop lijkende woordcombinaties worden gebruikt. Het begrip 'antwoordtendentie' of 'antwoordgedrag' lijkt minder passend omdat hieronder ook, of vooral, verandering van de 'ware' waarde wordt verstaan. Antwoordstijl ziet men terug bij antwoorden op vragen van allerlei aard, die dezelfde schaal gebruiken, onafhankelijk van het onderwerp.

perspectief worden gezet. Vervolgens kunnen alle antwoorden individueel worden herschaald (recalibratie, correctie) naar een gemeenschappelijke schaal, wat weer vergelijkingen binnen en tussen landen mogelijk maakt die veel minder vertekend zijn.

In dit hoofdstuk gebruikten wij de WHS gegevens van ongeveer 150,000 respondenten in 64 landen, die steeds 5 zgn. calibratie vignetten bevatten voor elk van de 8 domeinen. Met multi-level ordered probit regressie kon antwoordstijl-variatie in relatie tot land en persoonlijke kenmerken in beeld worden gebracht. Het multilevel model ging uit van 2 niveaus (levels), individu (persoonlijke kenmerken) en land (nationale kenmerken); het stelde antwoordstijl vast per land (b.v. veroorzaakt door maatschappelijke normen) en gerelateerd aan het individuele niveau (b.v. gerelateerd aan opleiding). Landgebonden patronen (boven-individueel) werden beschreven met de intraclass correlatie coëfficiënt per vignet en domein, terwijl individuele antwoordstijl patronen beschreven werden in termen van a) lokale samentrekking of verwijding, en b) op-, dan wel neerwaartse verschuiving (shift) van de schaal. Landgebonden patronen in antwoordstijl waren vooral belangrijk bij de domeinen respect voor privacy, zorgtoegankelijkheid en kwaliteit voorzieningen.

Individu-gebonden patronen zagen we op 2 manieren. Ten eerste was samentrekking/verwijding van de schaal afhankelijk van sexe, opleidingsniveau, het hebben van zorgverantwoordelijkheid voor een naaste met een chronische ziekte, en met het belang dat aan responsiveness als gezondheidssysteem criterium werd toegekend. Ten tweede werd een verschuivingseffect gevonden in relatie tot eigen gezondheid, waarbij een slechte eigen gezondheid systematisch een lagere responsiveness score geeft (wat er bij correctie toe leidt dat de scores van de eigen ervaren responsiveness dus wat omhoog worden bijgesteld). Een sleutelbevinding was dat patronen van antwoordstijl dezelfde waren over de 8 responsiveness domeinen. Dit hoofdstuk liet zien dat antwoordstijl veel sterker en consistentier met individuele kenmerken samenhangt dan eerder gedacht, met overigens nog steeds aanzienlijke landgebonden effecten. Omdat deze kenmerken niet willekeurig zijn verdeeld, toont dit de noodzaak aan om voor antwoordstijl te corrigeren bij responsiveness vergelijkingen (en bij andere uitkomsten) wat vooral bij ongelijkheidsonderzoek belangrijk is. Het is daarbij gelukkig zo dat vignet antwoorden beïnvloed worden door maar een beperkt aantal tevoren veronderstelde factoren, die ook konden worden gemeten, waardoor correctie mogelijk was.

Universeel geldige determinanten van responsiveness: hoofdstukken 5-7

Hoofdstuk 5 onderzoekt of personen verschillend belang hechten aan de 8 responsiveness domeinen, in de wetenschap dat het aanvankelijke uitgangspunt van de vragenlijst was dat de 8 domeinen even belangrijk zijn, wereldwijd. Om dit na te gaan bevatte de MCS items die respon-

denten vroegen de responsivens domeinen te rangordenen naar belang. Gegevens afkomstig van 105,806 respondenten van de MCS in 41 landen werden gebruikt. Multinomiale logit regressie modellen werden toegepast om de vragen naar de belangrijkste en minst belangrijke domeinen te analyseren. Variatie in toegekend domein belang werd gerelateerd aan landgebonden kenmerken (land waar men woont, ontwikkelingsniveau, uitgaven aan gezondheidszorg, en een zgn. 'geographische zone' variabele) en aan individuele kenmerken (sexe, leeftijd, opleidingsniveau, eigen gezondheid, en zorggebruik (klinische zorg, ambulante zorg)). Doorgaans wezen respondenten zorgtoegankelijkheid aan als het belangrijkste domein. Bejegening was vervolgens het belangrijkste, gevolgd door communicatie. Sociale steun was het minst belangrijk. Er was consistentie tussen de antwoorden op de vragen naar de belangrijkste en minst belangrijke domeinen (na omkering van de volgorde uiteraard), met als uitzondering dat bejegening en zorgtoegankelijkheid van plaats wisselden. De rangorde consistentie was sterker binnen landen dan ertussen. De sterkste twee individuele factoren die de rangorde bepaalde waren eigen gezondheid en zorggebruik, waarbij gezondere personen en zij die alleen ambulante zorg gebruikten eerder bejegening, autonomie en keuzevrijheid als belangrijkste domein kozen, relatief t.o.v. zorgtoegankelijkheid. Wie ziek is wil allereerst geholpen worden. The landgebonden rangorde-effecten waren meest zichtbaar voor landen met een lagere ontwikkeling; in deze landen werd een wat hoger belang gehecht aan de kwaliteit van voorzieningen. Landen buiten de EU⁶ hechten duidelijk meer belang aan de mogelijkheid tot sociale steun (tijdens bv. ziekenhuisopname). Keuzevrijheid was relatief belangrijk in Centraal en Oost-Europa en landen van de voormalige Sovjet-Unie (geografische zone-indeling van de WHO); respect voor privacy werd hoger aangeslagen in het Midden Oosten, de EU, en landen van de voormalige Sovjet-Unie. Niettegenstaande deze variaties, was er een duidelijke wereldwijde volgorde: zorgtoegankelijkheid, bejegening en communicatie zijn het belangrijkste. Deze rangorde betekent niet dat de andere domeinen irrelevant zijn, alleen dat er desgevraagd een rangorde is. Deze bevinding geeft wel een suggestie hoe het aantal items verminderd kan worden, en waar - als het moet - bij kwaliteit van zorgverbetering de eerste aandacht voor moet zijn; ook bij vergelijking tussen landen kan langs deze volgorde geprioriteerd worden. Het belangrijkste resultaat van de vele onderliggende analyses was wellicht de wetenschappelijk steun voor een universele interpretatie van het responsivens concept.

Hoofdstuk 6 stelt de vraag of bepaalde kenmerken van een gezondheidszorgsysteem of van de professionals wereldwijd geassocieerd zijn met betere of minder goede responsivens. Dit is een politiek gevoelige vraag, vooral als men kan aannemen dat - zie voorgaande hoofdstukken - de vergelijking van landen niet vertekend wordt door antwoordstijl-effecten. De WHS bevatte een brede gegevensset over kenmerken van het gezondheidszorgsysteem (12 variabelen) ge-

6 Voor een exacte beschrijving van de regionale zone-indeling, zie de betreffende hoofdstukken.

meten op individueel niveau, en daarnaast 8 individu-gebonden determinanten; deze werden gebruikt in multilevel regressie analyses, voor elke combinatie van een domein met klinische dan wel ambulante responsiveness ervaringen. De WHS gegevensset in onze analyse bevatte ongeveer 120,000 respondenten uit 49 landen. Voor elk responsiveness domein werd op individueel niveau een dichotome score gemaakt van de oorspronkelijke 5 ordinale responsies (slechte responsiveness = “moderate”, “bad” of “very bad”). Hiermee kan een frequentie van slechte (of goede) responsiveness voor een groep worden berekend en kunnen verkennende regressie-analyses worden uitgevoerd. Het meest opvallende resultaat was de aanwezigheid van een bepaalde groep gezondheidszorgkenmerken op globaal niveau de gemeten responsiveness verklaarden, veel sterker dan de opgenomen individu-gebonden kenmerken. De volgende zorgkenmerken waren het belangrijkste en verminderde de responsiveness (doorgaans over verschillende domeinen): discriminerende houding van gezondheidsprofessionals, ontoegankelijkheid van zorgfaciliteiten in termen van gemeten reistijd naar de voorziening, en het gerund worden door de overheid in plaats van door niet-overheidsorganisaties. Bij individu-gebonden kenmerken viel op dat armoe en ongezonde eigen gezondheid, sterker dan leeftijd en sexe (wat doorgaans wordt aangenomen), van belang waren. Dit is in lijn met eerdere uitkomsten rond ongelijkheid (hoofdstuk 3). Deze en andere bevindingen tonen aan dat responsiveness een uitkomstmaat is die gevoelig is voor ongelijkheid. Aanvullende resultaten suggereren langs welke specifieke wegen dit effect in bepaalde domeinen tot stand komt.

Hoofdstuk 7 presenteert een ecologische regressie analyse op landniveau. Determinanten en uitkomsten zijn op landniveau gemeten (b.v. landelijke dekking van vaccinaties, kindersterfte). De gekozen aanpak veronderstelt dat responsiveness instrumenteel is (een zgn. intermediaire of mediërende variabele) tussen zorgtoegankelijkheid in brede zin en gezondheidsuitkomsten (zoals dekking van preventieve diensten en geaggregeerde gezondheidsuitkomsten). Verschillende regressiemodellen (negatief binomiaal, log-lineair and kleinste kwadraten) werden vergeleken om het beste model te vinden bij deze typen variabelen. Het bleek dat bejegening en zorgtoegankelijkheid geassocieerd waren met moedersterfte en sterfte aan tuberculose; bejegening alleen was significant voor kindersterfte. Hoewel soms effecten tussen responsiveness en gezondheid in beide richtingen gaan, gaan we meestal uit van een effect van responsiveness op uitkomst. Een sterk voorbeeld van het laatste lijkt de waargenomen ongelijke mazelenvaccinatie-dekking die afhangt van gemeten betaalbaarheid en gemeten responsiveness. Dit voorbeeld wijst op de relevantie van zowel betaalbaarheid als aanvaardbaarheid van het zorgaanbiedingsproces, om gelijke toegang tot interventie voor arm en rijk te realiseren. Deze sleutelbevindingen bij een analyse op landniveau zijn in lijn met vergelijkbare individu-gebonden associaties (bv. met inkomen), wat de kans op een zgn. ‘ecological fallacy’ kleiner maakt. Meer algemeen ondersteunt zo’n voorbeeld van een relevant effect van zowel responsiveness als betaalbaarheid de

validiteit van het "Health Systems Performance Assessment framework" van de WHO uit het jaar 2000.

Responsiveness meting toegepast in perinatale zorg: hoofdstukken 8-9

De laatste 2 hoofdstukken beschrijven de aanpassing van het responsiveness concept en de items gebruikt in de surveys naar de specifieke context van geboortezorg. De Nederlandse perinatale responsiveness vragenlijst zoals hier beschreven (ReproQ) werd gemaakt als samenwerking met een onderzoeksteam van ErasmusMC, Rotterdam en het UMC Utrecht. Voor de interviewer-ondersteunde ReproQ zoals hier beschreven, pasten we de WHS and MCS responsiveness items aan voor drie onderscheiden fases van geboortezorg: antenataal (25 items), geboorte (40 items), kraambed (39 items); de vragenlijst vroeg ook naar het belang van domeinen; ook werden kenmerken van het zorgproces gemeten (bv. verwijzing/discontinuïteit van zorg).

Hoofdstuk 8 beschrijft de psychometrisch kenmerken (haalbaarheid, betrouwbaarheid, construct en discriminatieve validiteit), in een steekproef van 171 Nederlandse pasbevallen vrouwen. Het face-to-face interview vond 2 weken post partum plaats (duur interview: tussen 20 en 40 minuten). Het percentage missende items was 8%. De gemiddelde Cronbach's alpha voor de antenatale, geboorte en kraambed items was respectievelijk 0.73, 0.84, en 0.87. Binnen de groep antenatale items, geboorte items en kraambed items verklaarden de 8 responsiveness domeinen 69%, 69%, and 76% van de variantie, een goed resultaat. Globale responsiveness scores waren hoger (beter) als geen ziekenhuisverwijzing had plaatsgevonden. Deze bevinding bevestigde de discriminatieve validiteit and de hypothese dat ontevredenheid met de gezondheidsuitkomsten van invloed is op het oordeel over de responsiveness van de ondervonden geboortezorg. Globaal gezien voldeed de interviewer-ondersteunde ReproQ wat betreft psychometrische kenmerken om verschillen in kwaliteit van geboortezorg aan vast te stellen.

Hoofdstuk 9 toont een groot aantal responsiveness resultaten, zoals verzameld met de hiervoor beschreven ReproQ. De nadruk lag op antenatale en geboorte resultaten. Responsiveness antwoorden werden net als in hoofdstuk 6, gedichotomiseerd (slechte responsiveness = "moderate", "bad" of "very bad"). Het toegekende domein belang in deze specifieke situatie kwam overeen met de rangorde gerapporteerd in de MCS en WHS (hoofdstuk 5). De frequentie van slechte responsiveness ervaringen varieerde van 5.9% tot 31.7% voor de antenatale items, en van 9.7% tot 27.1% voor de geboorte items. Zowel antenataal als tijdens de geboorte, scoorden de persoonlijke domeinen (b.v. bejegening en communicatie) beter dan 'setting' domeinen (b.v. zorgtoegankelijkheid, kwaliteit van voorzieningen). Een groep van onafhankelijke factoren gerelateerd aan zorgproces en persoonlijke achtergrond hadden invloed op responsiveness. Al

met al werd responsiveness meer verklaard door gezondheid- en gezondheidszorggerelateerde factoren dan door persoonlijke factoren, iets dat we ook eerder vonden bij de MCS en WHS (hoofdstuk 6). Zgn. 'known group comparisons' ondersteunden de validiteit van de ReproQ: zoals tevoren verondersteld, beïnvloedden obstetrische voorgeschiedenis en een slechte geboorteuitskomst de responsiveness uitkomsten in negatieve zin. De resultaten suggereren dat het oorspronkelijke responsiveness concept met bijbehorende vragen en meetmethodieken uitstekend kan worden aangepast aan specifieke klinische situaties of settings.

DWARSVERBANDEN - DISCUSSIE

Na de eerste presentatie van responsiveness in het WHO World Health Report 2000, werd de relevantie van responsiveness door sommige critici ter discussie gesteld. Op dat moment was er zo goed als geen psychometrische informatie, zoals beschreven in dit proefschrift. Onze resultaten legitimeren responsiveness als afzonderlijk en onmisbaar doel, naast gezondheid en financiële rechtvaardigheid, als de prestaties van een gezondheidszorgsysteem worden gekwantificeerd. Wij beschouwen het bewijs overtuigend ten gunste van het belang en de universaliteit van responsiveness, waarmee kwaliteit van het zorgproces op het individuele niveau kan worden gemeten. Het veronderstelde instrumentele belang van responsiveness voor gezondheidsuitkomsten, en het belang voor 'universal health coverage'⁷ en gelijkheid van gezondheid over landen werd aangetoond.

Responsiveness werd in de discussie over het meten van de kwaliteit van gezondheidszorgsystemen geïntroduceerd in 1999-2000; het dankt veel aan het oudere begrip 'patiënttevredenheid'. Er zijn echter subtiele maar belangrijke verschillen tussen responsiveness en patiënttevredenheid, meer op het gebied van responseformulering dan op het gebied van domeinkeuzes en daarvan afgeleide items. Het antwoord op een responsiveness vraag betreft altijd de mate waarin een ervaring feitelijk aanwezig of afwezig was; daarentegen refereert 'tevredenheid' aan een interne toestand of emotie. Ontevredenheid kan wijzen op veel verschillende achtergronden. Naast gebrek aan kwaliteit van zorg, kan het wijzen op het niet hebben van de ervaring als zodanig terwijl deze wel verwacht werd of nodig was. Deze meervoudige interpreterbaarheid van een slecht resultaat is een groot nadeel bij gebruik van patiënttevredenheid in kwaliteitsprocessen. Bij het responsiveness concept is de interpretatie van een response of van een somscore rechttoe-rechtaan, en verwachtingen spelen gezien de feitelijke vraag een veel kleinere rol, en voorzover die zich in antwoordstijl vertalen zijn ze te corrigeren. De thans waargenomen sterke toename in gebruik van responsiveness-achtige meetinstrumenten, aangeduid als 'patient re-

⁷ Dit is een specifiek begrip afkomstig van de WHO dat kortweg aanduidt dat in het optimale geval ieder individu verzekerd is van voldoende zorg, wat nog niet betekent dat ieder individu is verzekerd voor zorg.

ported experience measures' (PREMs), bv. in klinische kwaliteitsregistraties, binnen alle specialismen, bewijst overtuigend de universele relevantie van het concept.

Het meten van responsiveness via cliënt vragenlijsten lijkt robuust, geschikt voor verschillende settings en bestand tegen allerlei beperkingen. Bepaalde beschikbare methoden kunnen eventueel de kosten van toepassing verminderen, en andere methoden kunnen de sensitiviteit van de vragenlijst verhogen. Dit proefschrift toont aan dat het zelfs mogelijk is nuttige resultaten te verkrijgen als de responsiveness domeinen worden gereduceerd van 8 naar 4 sleuteldomeinen: zorgtoegankelijkheid, bejegening, communicatie en kwaliteit van de voorzieningen.

Ongeacht dit aantal domeinen, is het altijd gewenst vignetten toe te voegen om te kunnen corrigeren voor antwoordstijl. Misschien minder vignetten dan wij toepasten, nu we weten dat (individuele) antwoordstijl hetzelfde is voor verschillende domeinen. De sensitiviteit van de vragenlijst kan worden verhoogd door de volgende items toe te voegen aan de items van de 4 sleuteldomeinen: items over gemak van toegang bij zorgtoegankelijkheid, items over houding en gedrag bij bejegening, items over communicatie van professionals onderling bij communicatie, en de mate van ervaren georganiseerd zijn (logistieke kwaliteit) bij kwaliteit van voorzieningen.

AANBEVELINGEN

De algemene aanbeveling is dat elke vorm van zorgverlening op routinebasis ook de responsiveness moet meten. Deze aanbeveling gaat uit van het nut en de relevantie van continue kwaliteitsverbetering. Internationale organisaties zouden zich moeten inspannen om een standaard set van responsiveness items vast te stellen in veel gebruikte meetinstrumenten, vooral degenen die nu gebruikt worden door de WHO bij het meten van voortgang in het bereiken van Universal Health Coverage, het derde hoofddoel van de Sustainable Development Agenda die richting geeft aan WHO werkzaamheden, zoals andere doelen doen voor andere instellingen van de Verenigde Naties. Responsiveness resultaten moeten gecorrigeerd worden voor antwoordstijlen, tenzij uit de data blijkt dat dat niet nodig is. Binnen landen, moeten de vele klinische kwaliteitsregistraties de toepassing van vignet-gebaseerde correctie overwegen.

In kwaliteitsrapportages kan, als het gaat om simpele kwaliteitsdrempels en nationale benchmarks, worden volstaan met geaggregeerde responsiveness totaalscores (over alle domeinen). Maar alleen met totaalscores werken bevelen we niet aan, omdat elk domein afzonderlijke informatie bevat, vaak ook over specifieke causale factoren die van belang zijn als je de domeinkwa-

liteit gericht wil verbeteren. Correctie voor verschillen in antwoordstijl is in deze context meestal noodzakelijk, wat normaliter expliciete aandacht en wat extra middelen vereist.

Gezien het gemak waarmee dat mogelijk blijkt, bevelen we contextuele aanpassing van de responsiveness vragenlijst aan voor specifieke gezondheidszorg sub-systemen, gedefinieerd per medisch specialisme. Dergelijke aanpassingen zagen we al voor chronische reuma en de geboortezorg (de laatste niet alleen in Nederland maar op iets andere wijze ook in de UK 2015 maternity survey). Beroepsverenigingen, medische professionals, en managers van medische instellingen zouden moeten onderzoeken, ondersteund door responsiveness meting en de meting van procesvariabelen en uitkomsten, wat de meest effectieve kwaliteitscirkel is om responsiveness te verbeteren. Onderzoekers zouden zich moeten inspannen om harmonisatie tot stand te brengen tussen sets van generieke vragenlijsten (PREMs, PROMs); toekomstig onderzoek zou zich moeten richten op vereenvoudiging van de vignetten procedure; tenslotte zou ook de samenhang tussen responsiveness en toegankelijkheid maar ook therapietrouw onderzocht moeten worden, vooral voor achterstandsgroepen.

FULL BIBLIOGRAPHY



BIBLIOGRAPHY

- Aday L, Andersen R. A framework for the study of access to medical care. *Health Serv Res.* 1974;9:208–20.
- Agency for Healthcare Research and Quality. 2014 National healthcare quality and disparities report. Rockville MD: Department of health and human services, Agency for Healthcare Research and Quality; (May) 2015.
- Agency for Healthcare Research and Quality. CAHPS 2.0 survey and reporting kit. Rockville: Agency for Healthcare Research and Quality; 1999.
- Almeida C, Braveman P, Gold M, Szwarcwald CL, Ribeiro JM, Americo M. Methodological concerns and recommendations on policy consequences of the world health report 2000. *Lancet.* 2001; 257:1692–7.
- Anand S, Ammar W, Evans T, Hasegawa T, Kissimova-Skarbek K, Langer A, et al. Report of the scientific peer review group on health systems performance assessment. In: Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism.* Geneva: World Health Organization; 2003. p. 839–916.
- Andersen R. Revisiting the behavioural model and access to medical care: does it matter? *J Health Soc Behav.* 1995;36(1):1–10.
- Anhang Price R, Elliott MN, Zaslavsky AM, Hays RD, Lehrman WG, Rybowski L, et al. Examining the role of patient experience surveys in measuring health care quality. *Medical Care Res Rev.* 2014;71(5):522–54.
- Atkinson S. Political cultures, health systems and health policy. *Soc Sci Med.* 2002;55(1):113–24.
- Backhans M, Burström B, Ponce de Leon A, Marklund S. Is gender policy related to the gender gap in external cause and circulatory disease mortality? A mixed effects model of 22 OECD countries 1973-2008. *BMC Public Health.* 2012;12:e969. doi: 10.1186/1471-2458-12-969.
- Bago d'Uva T, van Doorslaer E, Lindeboom M. Does reporting heterogeneity bias the measurement of health disparities? *Health Econ.* 2008;17:351–75.
- Basu S, Andrews J, Kishore S, Panjabi R, Stuckler, D. Comparative performance of private and public health-care systems in low- and middle-income countries: a systematic review. *PLoS Med.* 9(6): p. e1001244.
- Benson T, Potts HWW. A short generic patient experience questionnaire: howRwe development and validation. *BMC Health Serv Res.* 2014;14:499.
- Berwick D, Nolan T, Whittington J. The Triple Aim: care, health, and cost. *Health Affair.* 2008;27(3):759–69.
- Berwick, D. A User's Manual For The IOM's 'Quality Chasm' Report. *Health Affair.* 2002; 21(3):80–90.
- Biemer PP, Lyberg LE. *Introduction to survey quality.* New York: John Wiley; 2003.
- Biró A. Copayments, gatekeeping, and the utilization of outpatient public and private care at age 50 and above in Europe. *Health Policy Plann.* 2013;111(1):24–33.
- Bleich S, Özaltın E, Murray CJL. How does satisfaction with the health-care system relate to patient experience? *Bull World Health Organ.* 2009;87:271–78.
- Blendon RJ, Kim M, Benson JM. The public versus the World Health Organization on health system performance. *Health Affair.* 2001;20(3):10–20.
- Blendon RJ, Schoen C, DesRoches CM, Osborn R, Zaper K. Common concerns amid diverse systems: health care experiences in five countries. *Health Affair.* 2003(May-Jun);22(3):106–21.

- Bodkera M, Pisingera C, Toft U, Jørgensena T. The rise and fall of the world's first fat tax. *Health Policy Plann.* 2015;119(6):697–850.
- Boehm K, Schmid A, Goetze R, Landwehr C, Rothgang H. Five types of OECD healthcare systems: empirical results of a deductive classification. *Health Policy.* 2013;113(3):258–69. doi: 10.1016/j.healthpol.2013.09.003.
- Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010. Rotterdam: Erasmus MC; 2010. [online: <http://www.nvk.nl/Nieuws/Dossiers/DossierPerinataleZorg.aspx>]
- Boshuizen H, Viet A, Picavet H, Botterweck A, van Loon A. Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. *Public Health.* 2006;20(4):297–308.
- Bradley E, McGraw S, Curry L, Buckser A, King K, Kasl S, et al. Expanding the Andersen model: the role of psychosocial factors in long-term care use. *Health Serv Res.* 2002;37(5):1221–42.
- Bramesfeld A, Wedegartner F, Elgeti H, Bisson S. How does mental health care perform in respect to service users' expectations? Evaluating inpatient and outpatient care in Germany with the WHO responsiveness concept. *BMC Health Serv Res.* 2007;7:e99.
- Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ.* 2002;21(2):271–92.
- Busse R, Valentine NB, Lessof S, Prasad A, van Ginneken E (2012). 'Being responsive to citizen's expectations: the role of health services in responsiveness and satisfaction'. In: Figueras J & McKee M (eds.), *Health Systems, Health, Wealth and Societal Well-being: assessing the case for investing in health systems*, Open University Press, Maidenhead. ISBN-13: 978-0-335-24430-0 (pb) ; ISBN-10: 0-335-24430-0; eISBN: 978-0-335-24431-7
- Cacace M, Ettelt S, Brereton L, Pedersen J, Nolte E. How health systems make available information on service providers. Experience in seven countries. Santa Monica: Rand, 2011. Accessed on 2 April 2016 from: http://www.rand.org/content/dam/rand/pubs/technical_reports/2010/RAND_TR847.pdf
- Campbell A, Converse PE, Rodgers WL. The quality of American life; perceptions evaluations and satisfactions. New York: Russell Sage Foundation; 1976.
- Campbell SM, Roland MO, Buetow SA. Defining quality of care. *Soc Sci Med.* 2000; 51(11):1611–25.
- Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). *Soc Sci Med.* 1997;44(5):681–92.
- Christiaens W, Gouwy A, Bracke P. Does a referral from home to hospital affect satisfaction with childbirth? A cross-national comparison. *BMC Health Serv Res.* 2007;7:e109. doi: 10.1186.
- Chuang Y, Sung P, Chao H, Bai C, Chang C. A longitudinal ecological study of the influences of political, economic, and health services characteristics on under-five mortality in less-developed countries *Health Place.* 2013;23 (September):111–21.
- Clark TS, Linzer DA. Should I use fixed or random effects? The society for political methodology. *Political Science Research and Methods.* 2015;3(2):399–408.
- Cleary PD, Edgman-Levitan S, Walker JD, Gerteis M, Delbanco TL. Using patient reports to improve medical care: a preliminary report from 10 hospitals. *Qual Manag Health Care.* 1993;2(1):31–8.
- Cleary PD. The increasing importance of patient surveys. *Brit Med J.* 1999; 319(720):721.

Commission of the European Communities. Together for health: a strategic approach for the EU 2008–2013. White paper. Com(2007) 630 final (Oct. 23, 2007). [online: http://ec.europa.eu/health/ph_overview/Documents/strategy_wp_en.pdf]

Commission on Social Determinants of Health. Final report. Closing the gap in a generation. Health equity through action on the social determinants of health. Geneva: World Health Organization; 2008.

Commission on the Measurement of Economic Performance and Social Progress. Report of the Commission on the Measurement of Economic Performance and Social Progress. Paris; 2008.

Costa-Font J, Zigante V. The choice agenda in European health systems: the role of 'middle class' demands. In: LEQS Paper No 82. London: London School of Economics and Political Science; (May) 2014.

Coulter A, Jenkinson C. European patients' views on the responsiveness of health systems and healthcare providers. *Eur J Public Health*. 2005;15(4):355–60.

Council of Australian Governments (GOAG). National Health Reform: performance and accountability framework. Canberra; 2015. Available from: <https://www.aihw.gov.au/getmedia/ea9b2361-38de-43f3-9426-8705fcc8f1da/performance-and-accountability-framework.pdf.aspx>

Couper M. Web surveys: a review of issues and approaches. *Public Opin Quart*. 2000; 64(4):464-94).

Cox DR, 2011 (August 23). Discussion on the paper by Rice, Robone and Smith (2012). In VS Isham (Chair), Vignettes and health systems responsiveness in cross-country comparative analyses. The 58th World Royal Statistical Society Annual Meeting, Dublin, Ireland.

Cramer D, Howitt D. The SAGE dictionary of statistics. London: Sage; 2004.

Crnkic K, Greenberg M, Ragozin A, Robinson N, Basham R. Effects of stress and social support on mothers and premature and full-term infants. *Child Dev*.1983;54(1):209–17.

de Leeuw E. Quality in mail, telephone, and face-to-face surveys. Amsterdam: TT-Publicaties; 1992.

de Leeuw E. To mix or not to mix data collection modes in surveys. *J Off Stat*. 2005; 21(2):233–55.

de Silva A. A framework for measuring responsiveness. Global programme on evidence for health policy discussion paper no. 32. Geneva: World Health Organization; 2000. Accessed on 2 April 2016 from <http://www.who.int/healthinfo/paper32.pdf>.

Delnoij DMJ, Rademakers JJ, Groenewegen PP. The Dutch Consumer Quality Index: an example of stakeholder involvement in indicator development. *BMC Health Serv Res*. 2010; 10:88–88.

Demott K, Bick D, Norman R, Ritchie G, Turnbull N, Adams C, Barry C, Byrom S, Elliman D, Marchant S, Mccandlish R, Mellows H, Neale C, Parkar M, Tait P, Taylor C. Clinical Guidelines and evidence review for post natal care: routine post natal care of recently delivered women and their babies. London: National Collaborating Centre for Primary Care (NICE) and Royal College of General Practitioners; 2006. [online: <http://www.nice.org.uk/nicemedia/live/10988/30146/30146.pdf>]

Department of Health: NHS. Outcomes Framework for England, 2013-2016. London: Department of Health, United Kingdom; (January, 23) 2013.

DerSarkissian M, Thompson CA, Arah OA. Time series analysis of maternal mortality in Africa. *J Epidemiol Commun H*. 2013;67:992–8.

DeVellis RF. Scale development. Theory and applications. London: Sage; 1991.

Dillman DA, Tarnai J. Mode effects of cognitively designed recall questions: a comparison of answers to telephone and mail surveys. In Biemer P, Groves RM, Lyberg LE, Mathiowetz NA, Sudman S, editors. New York:Wiley-Interscience Publication; 1991. p. 73–93.

- Dillman DA. Mail and internet surveys: the tailored design method. New York: John Wiley; 2000.
- Donabedian A. Aspects of medical care administration. Cambridge, MA: Harvard University Press; 1973.
- Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, MI: Health Administration Press; 1980.
- Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol*. 2006;59(10):1087–91.
- Dora C, Haines A, Balbus J, Fletcher E, Adair-Rohani H, Alabaster G. Indicators linking health and sustainability in the post-2015 development agenda. *Lancet*. 2014;385(9965):380–91.
- Dour H, Wiley J, Roy-Byrne P, Stein M, Greer Sullivan J, Sherbourne C, Bystritsky A, Rose R, Craske M. Perceived social support mediates anxiety and depressive symptom. Changes following primary care intervention. *Depress Anxiety*. 2013; 31(5):436–42.
- Edwards C, Staniszwesda S, Crichton N. Investigation of the ways in which patients' reports of their satisfaction with healthcare are constructed. *Sociol Health Ill*. 2004;26(2):159–83.
- Elliott MN, Edwards C, Hambarsoomians K, Angeles J, Hays RD. Patterns of unit and item nonresponse in the CAHPS Hospital Survey. *Health Serv Res*. 2005;40(6p2):2096–119.
- EURO-PERISTAT project in collaboration with SCPE EaE, 2008. European perinatal health report. Better Statistics for Better Health for Pregnant Women and Their Babies in 2004.
- Evans DB, Hsu J, Boerma T. Universal health coverage and universal access. *B World Health Organ*. 2013;91:546-A.
- Evans T, Brown H. Road traffic crashes: operationalizing equity in the context of health sector reform. *Int J Inj Control Sa*. 2003;10(1):11–12.
- Farley DO, Elliott MN, Haviland AM, Slaughter ME, Heller A. Understanding variations in Medicare consumer assessment of health care providers and systems scores: California as an example. *Health Serv Res*. 2011;46(5):1646–62.
- Fiorentini G, Ragazzi G, Robone S. Are bad health and pain making us grumpy? An empirical evaluation of reporting heterogeneity in rating health system responsiveness. *Soc Sci Med*. 2015;144 (November 2015):48–58. doi: 10.1016/j.socscimed.2015.09.009.
- Fischer E, Shumway M, Owen R. Priorities of consumers, providers and family members in the treatment of schizophrenia. *Psych Serv*. 2002;53(6):724–9.
- Frechtling J, Sharp L. The user-friendly handbook for mixed-method evaluations. Arlington, VA: Directorate for Education and Human Resources; 1997(August). Contract no.: 94-52965. Sponsored by the National Science Foundation. [online: www.nsf.gov/pubs/1997/nsf97153/start.htm]
- Fritzell J, Kangas O, Hertzman J, Blomgren J, Hiilamo H. Cross-temporal and cross-national poverty and mortality rates among developed countries. *J Environ Public Health*. 2011;2011:e915490. doi: 10.1155/2013/915490.
- Fung CH, Elliott MN, Hays RD, Kahn KL, Kanouse DE, McGlynn EA et al. Patients' preferences for technical versus interpersonal quality when selecting a primary care physician. *Health Serv Res*. 2005;40(4):957–77.
- Gakidou E, Cowling K, Lozano R, Murray CJL. Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis. *Lancet*. 2010;376:959–74.
- Gakidou E, Murray CJL, Evans DB. Quality and equity: preferences for health system outcomes. In Murray CJL, Evans DB, editors. *Health systems performance assessment*. Geneva:World Health Organization; 2003: 667–81.

- Gakidou E, Oza S, Vidal F, Li AY, Lee DK, Sousa A. Improving child survival through environmental and nutritional interventions: the importance of targeting interventions toward the poor. *JAMA*. 2007;298:1876–87.
- Gandek B, Ware JE, Jr., Aaronson NK, Alonso J, Apolone G, Bjorner J, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51(11):1149–58.
- Garratt A, Solheim E, Danielsen K. National and cross-national surveys of patient experiences: a structured review. Report no. 7-2008. Oslo: Norwegian Knowledge Centre for the Health Services; 2008.
- Gayet-Ageron A, Agoritsas T, Schiesari L, Kolly V, Perneger TV. Barriers to participation in a patient satisfaction survey: who are we missing? *PLoS One*. 2011;6(10):e26852.
- Gilson L, Alilio M, Heggenhougen K. Community satisfaction with primary health care services: an evaluation undertaken in the Morogoro region of Tanzania. *Soc Sci Med*. 1994;39(6):767–80.
- Gilson L, Doherty J, Loewenson R, Francis V. Challenging inequity through health systems: final report. Knowledge Network on Health Systems (WHO Commission on Social Determinants of Health). Geneva: World Health Organization; 2007. [online: http://www.who.int/social_determinants/resources/csdh_media/hskn_final_2007_en.pdf]
- Goddard M, Smith P. Equity of access to health care services: theory and evidence from the UK. *Soc Sci Med*. 2001;53:1149–62.
- Gostin L, Hodge J, Valentine N, Nygren-Krug H. The domains of health responsiveness: a human rights analysis. Health and human rights working paper no. 2. Geneva: World Health Organization; 2003. [online: http://apps.who.int/iris/bitstream/10665/73926/1/HHRWPS2_eng.pdf]
- Gouveia G, Luna C, Souza-Júnior PR, Szwarcwald CL. Health care users' satisfaction in Brazil. *Cad Saude Publica*. 2005;21:109–18.
- Greener I. Towards a history of choice in UK health policy. *Sociol Health Ill*. 2009; 31(3):309–24.
- Groenewegen PP, Kerssens JJ, Sixma HJ, van der Eijk I, Boerma WGW. What is important in evaluating health care quality? An international comparison of user views. *BMC Health Serv Res*. 2005;5(16):1–9.
- Grogan S, Conner M, Norman P, Willits D, Porter I. Validation of a questionnaire measuring patient satisfaction with general practitioner services. *Qual Safe Health Care*. 2000;(9): 210–5.
- Grol R, Wensing M, Mainz J, Ferreira P, Hearnshaw H, Hjortdahl P et al. Patient priorities with respect to general practice care: an international comparison. European Task Force on Patient Evaluations of General Practice (EUROPEP). *Fam Pract*. 1999;15(1): 4–11.
- Grol-Prokopczyk H, Freese J, Hauser RM. Using anchoring vignettes to assess group differences in general self-rated health. *J Health Soc Behav*. 2011;52(2):246–61.
- Groot W. Adaptation and scale of reference bias in self-assessments of quality of life. *J Health Econ*. 2000;19:403–20.
- Groves R. Survey error models and cognitive theories of response behaviour. In: Sirken M, Herrmann D, Schechter S, Schwartz N, Tanur J, Tourangeau R, editors. *Cognition and survey research*. New York: Wiley. p. 235–50.
- Gulliford M. Modernizing concepts of access and equity. *Health Econ Policy L*. 2009;4(pt2): 223–30. Doi:10.1017/S1744133109004940.

- Hajizadeha M, Heymann J, Strumpf E, Harper S, Nadi A. Paid maternity leave and childhood vaccination uptake: longitudinal evidence from 20 low-and-middle-income countries. *Soc Sci Med.* 2015 (September);140:104–17.
- Hakkinen U, Ollila E, Taipale V, Koivusalo U, Rissanen P, Sintonen H. The world health report 2000: what does it tell us about health systems? Analyses by Finnish experts in Themes from Finland 7/2000. Helsinki:National Research and Development, Center for Welfare and Health (STAKES); 2000.
- Hall J, Feldstein M, Fretwell M, Rowe J, Epstein A. Older patients' health status and satisfaction with medical care in an HMO population. *Med Care.* 1990;28:261–70.
- Hall JA, Dornan MC. Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis. *Soc Sci Med.* 1990;30:811–8.
- Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Satisfaction, gender, and communication in medical visits. *Med Care.* 1994;32(12):1216–1231.
- Hargraves JL, Hays RD, Cleary PD. Psychometric properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 adult core survey. *Health Serv Res.* 2003;38(6 Pt 1):1509–27.
- Hargraves JL, Wilson IB, Zaslavsky A, James C, Walker JD, Rogers G, et al. Adjusting for patient characteristics when analysing reports from patients about hospital care. *Med Care.* 2001;39(6):635–41.
- Harper S, Lynch J. Measuring health inequalities. In: Oakes J, Kaufman J, editors. *Methods in social epidemiology.* San Francisco: John Wiley & sons; 2006. p. 134–68.
- Health System Change. An update on the community tracking study: a focus on the changing health system. Issue Brief. Centre for Studying Health System Change, Mathematica Policy Research. Issue brief no 18. 1999(Feb). [online: www.hschange.org/content/63/63.pdf]
- Heinzman C, Kaplan R. Assessment of methods for measuring social support. *Health Psychol.* 1088;7(1):75–109.
- Holzer BM, Minder CE. A simple approach to fairer hospital benchmarking using patient experience data. *Int J Qual Health Care.* 2011;23(5):524–30.
- Hosseinpoor AR, Williams JA, Itani L, Chatterji S. Socioeconomic inequality in domains of health: results from the World Health Surveys. *BMC Public Health.* 2012; 12(198). doi: 10.1186/1471-2458-12-912.
- Hosseinpoor AR, Bergen N, Schlotheuber A. Promoting health equity: WHO health inequality monitoring at global and national levels. *Glob Health Action.* 2015(Sep);8e29034.doi: 10.3402/gha.v8.29034.
- Hudson S, Weisman C, Anderson R, Camacho F. The development and validation of the primary care satisfaction survey for women. *Women Health Iss.* 2004(Mar-Apr);14(2):35–50.
- Independent Expert Advisory Group (IEAG) on a Data Revolution for Sustainable Development. A world that counts. Mobilising the data revolution for sustainable development. New York:United Nations; 2014 (November).
- Inglehart R. Globalization and postmodern values. *Wash Q.* 2000(winter);23(1):215–28.
- Jasso G. Factorial survey methods for studying beliefs and judgments. *Sociol Method Res.* 2006 February;34(3):334–423.
- Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Health Care.* 2002;14(5):353–8.
- Jung HP, Baerveldt C, Olesen F, Grol R, Wensing M. Patient characteristics as predictors of primary health care preferences: a systematic literature analysis. *Health Expect.* 2003;6(2):160–81.

- Jurges H, van Soest A. Comparing the well-being of older Europeans: introduction. *Soc Indic Res.* 2012, May;105:187–90.
- Kaufman J, Oakes JE. *Methods in Social Epidemiology*. 2nd ed. John Wiley & Sons; 2017.
- Kelley E, Hurst J. Health care quality indicators project conceptual framework paper. Organisation for Economic Co-operation and Development (OECD) Working Papers 2006. Paris: OECD; 2006.
- Kerssens JJ, Groenewegen PP, Sixma HJ, Boerma WGW, van der Eijk I. Comparison of patient evaluations of health care quality in relation to WHO measures of achievement in 12 European countries. *B World Health Organ.* 2004;82(2):106–14.
- Kim J, Mueller CW. Factor analysis statistical methods and practical issues. In: Uslander EM, editor. *Quantitative applications in the social sciences*. London: Sage; 1978.
- King G, Murray CJL, Salomon J, Tandon A. Enhancing the validity and cross-cultural comparability of measurement in survey research. *Am Polit Sci Rev.* 2004 (February);98(1):191–207.
- King N, Harper S, Young M. Who cares about health inequalities? Cross-country evidence from the World Health Survey. *Health Policy Plann.* 2013;28(5):558–71.
- Kok R, Avendano Pabon M, Bago d'Uva T, Mackenbach J. Can reporting heterogeneity explain differences in depressive symptoms across Europe? *Soc Indic Res.* 2012;105:187–90.
- Kolves K, Milner A, Varn P. Suicide rates and socioeconomic factors in Eastern European countries after the collapse of the Soviet Union: trends between 1990 and 2008. *Sociol Health Ill.* 2013;35(6):956–70.
- Kowal P, Naidoo N, Williams S, Chatterji S. Performance of the health system in China and Asia as measured by responsiveness. *Health.* 2011;3(10):638–45.
- Krol MW, de Boer D, Rademakers JJ, Delnoij DM. Overall scores as an alternative to global ratings in patient experience surveys: a comparison of four methods. *BMC Health Serv Res.* 2013;13:479.
- Kunrintemi V, Spatz ES, Di Capua P, et al. Patient–Provider communication and health outcomes among individuals with atherosclerotic cardiovascular disease in the United States. *Circ Cardiovasc Qual Outcomes.* 2017;10(4):e003635 (8 p.).
- Labarere J, Francois P, Auquier P, Robert C, Fourny M. Development of a French inpatient satisfaction questionnaire. *Int J Qual Health C.* 2001;13:99–108.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–74.
- Liabsuetrakul T, Petmanee P, Sanguanchua S, Oumudee N. Health system responsiveness for delivery care in Southern Thailand. *Int J Qual Health Care.* 2012;24(2):169–75.
- Lindeboom M, van Doorslaer E. Cut-point shift and index shift in self-reported health. *J Health Econ.* 2004;23:1083–99.
- Lopez M, Cooper L. Social support measures review. Final report. National Centre for Latino Child and Family Research; Laytonsville MD. (February 20) 2011. Accessed on 6 April 2018 from: https://www.first5la.org/files/SSMS_LopezCooper_LiteratureReviewandTable_02212011.pdf
- Lumley J. Assessing satisfaction with childbirth. *Birth.* 1985;12(3):141–5.
- Luo Q, Wang Q, Lu Z, Liu J. Evaluation of responsiveness of community health services in urban China: a quantitative study in Wuhan City. *PLoS One*;8(5):e62923.

- Mackenbach JP, Hu Y, Looma CW. Democratization and life expectancy in Europe, 1960–2008. *Soc Sci Med*. 2013;93:166–75.
- Malhotra C, Kyung Do Y. Socio-economic disparities in health system responsiveness in India. *Health Policy Plann*. 2013;28(2):197–205.
- Marmor T, Wendt C. Conceptual frameworks for comparing healthcare politics and policy. *Health Policy Plann*. 2012;107:11–20.
- McKee M. The world health report 2000: 10 years on. *Health Policy Plann*. 2010;25(5):346–8.
- McPake B. User charges for health services in developing countries: a review of the economic literature. *Soc Sci Med*. 1993;36(11):1397–405.
- Mirzoev T, Kane S. What is health systems responsiveness? Review of existing knowledge and proposed conceptual framework. *BMJ Global Health* 2017;2(4):e000486.
- Mohangoo AD, Buitendijk SE, Hukkelhoven CW, Ravelli AC, Rijninks-van Driel GC, Tamminga P, Nijhuis JG. [Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study]. *Ned Tijdschr Geneesk*. 2008;152(50):2718–27.
- Mondal NI, Shitan M. Relative importance of demographic, socioeconomic and health factors on life expectancy in low- and lower-middle-income countries. *J Epidemiol*. 2014;24(2):117–24.
- Montagu D, Yarney G, Visconti A, Harding A, Yoong J. Where do poor women in developing countries give birth? A multi-country analysis of demographic and health survey data. *PLoS One*. 2001;6(2). doi: 10.1371/journal.pone.0017155.
- Mooney G, Wiseman V. World health report 2000: challenging a world view. *J Health Res Policy Syst*. 2000;5:198–9.
- Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet*. 2007;370:851–8.
- Muldoon KA, Galway LP, Nakajima M, Kanters S, Hogg R, Bendavid E, et al. Health system determinants of infant, child and maternal mortality: a cross-sectional study of UN member countries. *Global*. 2011;7(42):e1–10.
- Murphy-Cullen CL, Larsen LC. Interaction between the socio-demographic variables of physicians and their patients: its impact upon patient satisfaction. *Soc Sci Med*. 1984;19(2):163–6.
- Murray CJL, Frenk J. A framework for assessing the performance of health systems. *Bull World Health Organ*. 2000;78(6):717–31.
- Murray CJL, Kawabata K, Valentine NB. People's experience versus people's expectations. *Health Affair*. 2001;20(3):41–3.
- National Institute for Public Health and Environment. Dutch health care performance report 2014 (in Dutch: *Zorgbalans 2014. De prestaties van de Nederlandse gezondheidszorg*). In. Edited by van den Berg MJ, De Boer D, Gijsen R, Heijink R, Limburg L, Zwakhals S. Bilthoven: National Institute for Public Health and Environment (RIVM); 2015.
- Netherlands Ministry of Housing, Spatial Planning and the Environment (in Dutch: *VROM*). *Stad en Stijging. Sociale stijging als leidraad voor stedelijke vernieuwing. [City and Rise. Social advancement as a guideline for urban renewal]*. The Hague. Publication advice No. 054. 2006.
- NHS Outcomes Framework Reports. London: Health and Social Care Information Centre; (February) 2016. Accessed on 6 April from: <https://digital.nhs.uk/catalogue/PUB30226>

Nilsson E, Orwelius L, Kristenson M. Patient-reported outcomes in the Swedish National Quality Registers. *J Intern Med* 2015 277(1):141–53.

Nunnally JCB, Bernstein IH. *Psychometric theory*. 3rd ed. New York: McGraw-Hill; 1994.

Onukwughu E, Saunders E, Daniel Mullins C, Pradel FG, Zuck M. A qualitative study to identify reasons for discharges against medical advice in the cardiovascular setting. *BMJ Open*, 2012;2:e000902. doi:10.1136/bmjopen-2012-000902.

Orly M, Matthews S, Power C. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *Int J Epidemiol*. 2000;29(1):149–57.

Ortiz JP, Valentine NB, Gakidou E, Tandon A, Kawabata K, Evans DB, Murray CJL. Inequality in responsiveness: population surveys from 16 OECD countries. In: Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism*. Geneva: World Health Organization; 2003. p.653–65.

Osgood DW. Poisson-based regression analysis of aggregate crime counts. *J Quant Criminol*. 2000;16(1):21–43.

Pearson M, Colombo F, Murakami Y, James C. Universal health coverage and health outcomes. Final Report. OECD: Paris; 2016 (22 July). Accessed on 6 April 2018 from: <https://www.oecd.org/els/health-systems/Universal-Health-Coverage-and-Health-Outcomes-OECD-G7-Health-Ministerial-2016.pdf>

Peltzer K. Patient experiences and health system responsiveness in South Africa. *BMC Health Serv Res*. 2009;9(1):1–12. doi: 10.1186/1472-6963-9-117.

Penchansky R, Thomas JW. The concept of access. *Med Care*. 1981;19(2):127–40.

Peracchi F, Rossetti C. Heterogeneity in health responses and anchoring vignettes. *Empir Econ*. 2012;42:513–38.

Peterson WE, Charles C, DiCenso A, Sword W. The Newcastle satisfaction with nursing scales: a valid measure of maternal satisfaction with inpatient postpartum nursing care. *J Adv Nurs*. 2005;52(6):672–81.

Pickard S, Sheaff R, Dowling B. Exit, voice, governance and user-responsiveness: the case of English primary care trusts. *Soc Sci Med*. 2006;63(2):373–83.

Picker Institute. *The Picker Institute implementation manual*. Boston, MA: Picker Institute; 1999.

Pillinger R. An advanced user's guide to Stat-JR version 1.0.0. Measuring dependency Bristol: University of Bristol; 2013 [cited 2014 February 5]. A transcript of measuring dependency presentation. [online: <http://www.bristol.ac.uk/cmm/learning/videos/measuring-dep.html#top>]

Poeran J, Valentine NB, Bonsel GJ, Zauber A, Moore J (2014). Statistics in Epidemiology: Communication is the Key to Success. Joint Statistical Meeting, 2–7 August 2014. Abstract available at: <http://www.amstat.org/meetings/jsm/2014/onlineprogram/AbstractDetails.cfm?abstractid=311354>

Rabe-Hesketh S, Skrondal A, Pickles A. *GLAMM Manual*. University of California Berkeley Division of Biostatistics Working Paper Series, Berkely: University of California, 2004. Accessed on 20 May 2016 from: <http://www.bepress.com/ucbbiostat/paper160>.

Rabe-Hesketh S, Skrondal A, Pickles A. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *J Econometrics*. 2005;128:301–23.

Rabe-Hesketh S, Skrondal A. *Multilevel and longitudinal modeling using Stata* (3rd ed.). College Station: Stata Press; 2012.

- Raleigh VS, Hussey D, Seccombe I, Hallt K. Ethnic and social inequalities in women's experience of maternity care in England: results of a national survey. *J R Soc Med.* 2010;103(5):188–98.
- Ravelli AC, Tromp M, van Huis M, Steegers EA, Tamminga P, Eskes M, Bonsel GJ. Decreasing perinatal mortality in The Netherlands, 2000–2006: a record linkage study. *J Epidemiol Commun H.* 2009;63(9):761–5.
- Redshaw M, Heikkila K. *Delivered with care: a national survey of women's experience of maternity care.* Oxford:National Perinatal Epidemiology Unit; 2010. [online: <https://www.npeu.ox.ac.uk/delivered-with-care>]
- Rice N, Robone S, Smith P. Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness. *Eur J Health Econ.* 2011;12:141–62.
- Rice N, Robone S, Smith P. The measurement and comparison of health system responsiveness. Health econometrics and data group (HEDG) working paper 08/05. York: University of York; 2008.
- Rice N, Robone S, Smith P. Vignettes and health systems responsiveness in cross-country comparative analyses. *J R Stat Soc A Stat.* 2012;175(2): 337–69.
- Rijnders M, Baston H, Schonbeck Y, van der Pal K, Prins M, Green J, Buitendijk S. Perinatal factors related to negative or positive recall of birth experience in women 3 years postpartum in the Netherlands. *Birth.* 2008;35(2):107–16.
- Robone S, Rice N, Smith P. Health systems' responsiveness and its characteristics: a cross-country comparative analysis. *Health Serv Res.* 2011;46(6pt2): 2079–100. doi: 10.1111/j.1475-6773.2011.01291.x.
- Rodgers G. Income and inequality as determinants of mortality: an international cross-section analysis. *Pop Stud-J Demog.* 1979;33(2):343–51.
- Ruger J, Kim H. Out-of-pocket healthcare spending by the poor and chronically ill in the Republic of Korea. *Am J Pub Hlth.* 2007;97(5):804–11.
- Russel S. Patients' experiences: top heavy with research. Literature Review. Melbourne: Research Matters; (June) 2013. Available from: <http://www.research-matters.com.au/publications/PatientsExperiencesReview.pdf>
- Russell E, Johnson B, Larsen H, Lelinneth M, Novilla B, van Olmen J, et al. Health systems in context: a systematic review of the integration of the social determinants of health within health systems frameworks. *Rev Panam Salud Publica.* 2013;34(6):461–7.
- Ryan M. Discrete choice experiments in health care. *Brit Med J.* 2004;328(7436):360–1.
- Salomon J, Tandon A, Murray CJL. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. *Brit Med J.* 2004;328(7434):258.
- Schafer JL. *NORM: multiple imputation of incomplete multivariate data under a normal model [statistical software].* University Park: Penn. State University. Department of Statistics; 2001.
- Scheerhagen M, van Stel H, Tholhuijsen D, et al. Applicability of the ReproQ client experiences questionnaire for quality improvement in maternity care. *PeerJ* 2016;4:e2092 (16 p).
- Sekimoto M, Asai A, Ohnishi M, Nishigaki E, Fukui T, Shimbo T et al. Patients' preferences for involvement in treatment decision-making in Japan. *BMC Fam Prac.* 2004(Mar);5(1)e1–10. doi:10.1186/1471-2296-5-1.
- Shengelia B, Tandon A, Adams O, Murray CJL. Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy. *Soc Sci Med.* 2005;61(1):97–109.
- Shiferaw S, Spigt M, Godefrooi M, Melkamu Y, Tekie M. Why do women prefer home births in Ethiopia? *BMC Preg Childbirth.* 2013;13:5.

Additional matter. Health System Responsiveness by N Valentine

- Sirven N, Santos-Eggimann B, Spagnoli J. Comparability of health care responsiveness in Europe. *Soc Indic Res.* 2012;105:255–71
- Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med.* 1997;45(12):1829–43.
- Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care.* 1998;10(4):311–7.
- Sitzia J. How valid and reliable are patient satisfaction data? An analysis of 195 studies. *Int J Qual Health Care.* 1999;11(4):319–28.
- Sixma HJ, van Campen C, Kerssens JJ, Peters L. Quality of care from the perspective of elderly people: the QUOTE-elderly instrument. *Age Aging.* 2000;29(2):173–8.
- Smedby B, Andersen R. International comparisons of health care systems: conceptual and methodological developments over half a century. *Soc Med Tidskr.* 2010;5-6:439–52.
- Smith LF. Development of a multidimensional labour satisfaction questionnaire: dimensions, validity, and internal reliability. *Qual Health Care.* 2001;10(1):17–22.
- Smith LF. Postnatal care: development of a psychometric multidimensional satisfaction questionnaire (the WOMBPN SQ) to assess women's views. *Br J Gen Pract* 2011, 61(591):e628–37.
- Smith P. Health system performance synthesis report. Luxembourg: Directorate-General Employment, Social Affairs and Inclusion, European Union; 2014.
- Sofaer S, Crofton E, Goldstein C, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? *Health Serv Res.* 40(6p2):2018–37.
- Solar O, Irwin A. A conceptual framework for action on the social determinants of health. Social determinants of health discussion paper 2. 2nd 2010 ed. Geneva: World Health Organization; 2007.
- Steine S, Finset A, Laerum E. A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. *Fam Pract.* 2001;18(4):410–9.
- Stewart M. Reflections on the doctor–patient relationship from evidence and experience. *Brit J Gen Pract.* 2005;55(519):793–801.
- Szwarcwald C, Souza-Júnior P, Damacena G. Socioeconomic inequalities in the use of outpatient services in Brazil according to health care need: evidence from the World Health Survey. *BMC Health Serv Res.* 2010;10(1):1–7.
- Tanahashi T. Health service coverage and its evaluation. *Bull World Health Organ.* 1978;56(2):295–303.
- Thompson SC, Pitts JS, Schwankovsky L. Preferences for involvement in medical decision-making: situational and demographic influences. *Patient Educ Couns.* 1993 (Dec 31);22(3):133–40.
- Truijens SE, Pommer AM, van Runnard Heimel PJ, Verhoeven CJ, Oei SG, Pop VJ. Development of the Pregnancy and Childbirth Questionnaire (PCQ): evaluating quality of care as perceived by women who recently gave birth. *Eur J Obstet Gynecol Reprod Biol.* 174(March 2014):35–40.
- United Nations Development Agency (UNDP). Human development report. UNDP: New York; 2001.
- United Nations Economic and Social Council: Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators. In: E/CB3/2016/2/Rev1. Edited by Commission S. New York: United Nations; 19 February 2016.
- United Nations Educational, Scientific and Cultural Organization [Internet]. Paris: UNESCO Data Centre, 'Education' statistics [cited May 2005]. Available from: <http://www.uis.unesco.org/>

United Nations General Assembly: Transforming our world: the 2030 Agenda for Sustainable Development (Resolution 70/1). New York, NY: United Nations; 2015.

United Nations Task Team for the Global Thematic Consultation on Health in the Post-2015 Development Agenda. Health in the post-2015 development agenda. Report of the global thematic consultation on health. The world we want. New York: United Nations; 2013.

Üstün TB, Chatterji S, Mechbal A, Murray CJL, WHS Collaborating Groups. The World Health Surveys. Chapter 58. In: Murray CJL, Evans DB, editors. Health Systems Performance Assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003. p.797–808.

Üstün TB, Chatterji S, Mechbal A, Murray CJL. Quality assurance in surveys: standards, guidelines and procedures. Household surveys in developing and transition countries: design, implementation and analysis. New York: United Nations; 2005.

Üstün TB, Chatterji S, Villanueva M, Bendib L, Celik C, Sadana R, Valentine NB, Ortiz J, Tandon A, Salomon J, Cao Y, Wan Jun X, Özalın E, Mathers C, Murray CJL. WHO Multi-Country Survey Study on Health and Responsiveness 2000-2001. Global programme on evidence for health policy discussion paper no. 37. Geneva: World Health Organization; 2001.

Üstün TB, Chatterji S, Villanueva M, Bendib L, Celik C, Sadana R, et al. WHO Multi-Country Survey Study on Health and Responsiveness 2000-2001. In: Murray CJL, Evans DB, editors. Health systems performance assessment. Geneva:World Health Organization; 2003. p. 761–96.

Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for understanding health and health services. *Ann Fam Med.* 2009;7(4): 357–63. doi: 10.1370/afm.983.

Valentine NB (2011, August). Contribution to the discussion presented at the Ordinary Meeting of the Royal Statistical Society, Dublin, Ireland. “Vignettes and health systems responsiveness in cross-country comparative analyses” by Nigel Rice and Peter Smith. Published in: *Journal of the Royal Statistical Society Series A (Statistics in Society)* 01/2012; 175(2):360–1. doi: <http://dx.doi.org/10.2307/23251886>.

Valentine NB, Bonsel GJ (2003, June). Responsiveness for inpatient and outpatient health services from population surveys in 32 countries: differences by level of GDP and the level of civil society liberty. Paper presented at the International Health Economics Association (IHEA) Conference 2003, 14-19 June, San Francisco, USA. [No web link available]

Valentine NB, Bonsel GJ, Murray CJL. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Qual Life Res.* 2007;16(7):1107–25.

Valentine NB, Bonsel, GJ. Exploring models for the roles of health systems' responsiveness and social determinants in explaining universal health coverage and health outcomes. *Glob Health Action.* 2016(Mar);Suppl 9:e29329.doi: 10.3402/gha.v8.29329.

Valentine NB, Bonsel, GJ. What explains users' reports on health system responsiveness? Exploring health-care service and personal characteristics from surveys in 49 countries [Manuscript submitted for publication].

Valentine NB, Darby C, Bonsel GJ. Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of “health systems responsiveness” in 41 countries. *Soc Sci Med.* 2008;66(9):1939–50.

Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB. Health system responsiveness: concepts, domains and measurement. In: Murray CJL, Evans DB, editors. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003. p. 573–96.

Additional matter. Health System Responsiveness by N Valentine

Valentine NB, Lavallée R, Liu B, Bonsel GJ, Murray CJL. Classical psychometric assessment of the responsiveness instrument in the WHO Multi-Country Survey Study on Health and Responsiveness 2000–2001. In: Murray CJL, Evans DB, editors. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003. p. 597–630.

Valentine NB, Ortiz JP, Tandon A, Kawabata K, Evans DB, Murray CJL. Patient experiences with health services: population surveys from 16 OECD countries. In: Murray CJL, Evans DB, editors. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization. p. 643–52.

Valentine NB, Prasad A, Rice N, Ribone S, Chatterji S. Health systems responsiveness: a measure of the acceptability of health-care processes and systems from the user's perspective. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, editors. Performance measurement for health system improvement: experiences, challenges and prospects. Cambridge: Cambridge University Press; 2010. p. 138–86.

Valentine NB, Prasad A. Sample reports for 2000-01 Multi-Country Survey Study on Health and Health Systems Responsiveness [Internet]. Geneva:World Health Organization; 2005 [cited 2016 July 20]. Available from: www.who.int/responsiveness/en/

Valentine NB, Salomon J. Weights for Responsiveness Domains. In: Murray CJL, Evans DB, editors. Health Systems Performance Assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003. p. 631–42.

Valentine NB, Verdes-Tennant E, Bonsel GJ. Health systems' responsiveness and reporting behaviour: multilevel analysis of the influence of individual-level factors in 64 countries. *Soc Sci Med.* 2015;138:152–60.

van Buuren S. Appendices. Flexible imputation of missing data. Boca Raton, FL.: Chapman & Hall/CRC; 2012. p. 263–7.

van Campen C, Sixma JJ, Keressens HJ, Peters L, Rasker JJ. Assessing patients' priorities and perceptions of the quality of health care: the development of the QUOTE rheumatic-patients instrument. *Brit J Rheumatol.*1998;37(4):362–8.

van Cranenburgh OD, Krol MW, Hendriks MCP, de Rie MA, Smets EMA, de Korte J, Sprangers MAG. Consumer Quality Index Chronic Skin Disease (CQI-CSD): a new instrument to measure quality of care from the patient's perspective. *Brit J Dermatol.* 2015;173(4):1032–40.

van der Kooy J, Valentine NB, Birnie E, Vujkovic M, de Graaf JP, Denktas S, Steegers EA, Bonsel GJ. Validity of a questionnaire measuring the World Health Organization concept of health system responsiveness with respect to perinatal services in the Dutch obstetric care system. *BMC Health Serv Res.* 2014;14(1):e622. doi: 10.1186/s12913-04-0622-1.

van der Velden J. Dutch: Een goed begin.Veilige zorg rond zwangerschap en geboorte [Internet]. Royal Dutch Organisation of Midwives (KNOV); 2009 [cited 2015 May 20]. Available from: http://www.knov.nl/docs/uploads/Advies_stg_zw_en_geb_Een_goed_beginpdf_spread-versie_voor_media.pdf

van Doorslaer E, Masseria C, Koolman X et al. for the OECD Health Equity Research Group. Inequalities in access to medical care by income in developed countries. *Can Med Assoc J.* 2006; 174(2):177–83.

van Teijlingen ER, Hundley V, Rennie AM, Graham W, Fitzmaurice A. Maternity satisfaction studies and their limitations: "What is, must still be best". *Birth.* 2003; 30(2):75–82.

Vega J, Frenz P. Integrating social determinants of health in the universal health coverage monitoring framework. *Rev Panam Salud Publica.* 2013;34(6):468–72.

VROM. Stad en Stijging. Sociale stijging als leidraad voor stedelijke vernieuwing. [City and Rise. Social advancement as a guideline for urban renewal]. The Hague. Publication advice No. 054. 2006.

- Vujkovic M, Birnie E, Bonsel GJ. Maternal postnatal care utilisation in Rotterdam 2008-2009. The Rotterdam Public Health Report. Rotterdam: GGD Rotterdam Rijnmond; 2011.
- Wait S, Nolte E. Public involvement policies in health: exploring their conceptual basis. *Health Econ Policy Law*. 2006;1(2):149-62.
- Waldenstrom U, Nilsson CA, Winbladh B. The Stockholm birth centre trial: maternal and infant outcome. *Br J Obstet Gynaecol*. 1997;104(4):410-8.
- Waldenstrom U, Rudman A, Hildingsson I. Intrapartum and postpartum care in Sweden: women's opinions and risk factors for not being satisfied. *Acta Obstet Gynecol Scand*. 2006;85(5):551-60.
- Ware J. Scales for measuring general health perceptions. *Health Serv Res*. 1976 Winter; 396-415.
- Ware JE, Hays RD. Methods for measuring patient satisfaction with specific medical encounters. *Med Care*. 1988;26(4):393-402.
- Ware JE, Snyder MK, Wright WR, Davies AR. Defining and measuring patient satisfaction with medical care. *Eval Prog Plann*. 1983;6(3-4):247-93.
- Welzel C, Inglehart R, Klingemann H. The theory of human development: a cross-cultural analysis. *Eur J Polit Res*. 2003;42:341-79.
- Wensing M, Elwyn G. Methods for incorporating patients' views in health care. *Brit Med J*. 2003;326(7394):877-9.
- Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. *Soc Sci Med*. 1998;47(10):1573-88.
- Westaway MS, Rheeder P, Van Zyl DG, Seager JR. Interpersonal and organizational dimensions of patient satisfaction: the moderating effects of health status. *Int J Qual Health C*. 2003;15(4):337-44.
- Wieggers TA, van der Zee J, Keirse MJ. Transfer from home to hospital: what is its effect on the experience of childbirth? *Birth*. 1998;25(1):19-24.
- Wieggers TA. The quality of maternity care services as experienced by women in the Netherlands. *BMC Preg Childbirth*. 2009(May 9);9:e18. doi: 10.1186/1472-6963-12-69.
- Williams A. Science or marketing at WHO: a commentary on World Health 2000. *Health Econ*. 2001;10:93-100.
- Williams S, Calnan M. Convergence and divergence: assessing criteria of consumer satisfaction across general practice, dental and hospital care settings. *Soc Sci Med*. 1991;33(6):707-16.
- Wiysonge CS, Uthman OA, Ndumbe PM, Hussey GD. Individual and Contextual Factors Associated with Low Childhood Immunisation Coverage in Sub-Saharan Africa: a multilevel analysis. *PLoS One*. 2012;7(5):e37905. doi: 10.1371/journal.pone.0037905.
- Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials*. 2004;1(4):368-76.
- World Health Organization, EQUINET-REACH TRUST Malawi, TARSC Zimbabwe. Monitoring equity in access to AIDS treatment programmes: a review of concepts, models, methods and indicators Geneva: WHO; 2010.
- World Health Organization, International Health Partnership Initiative. Monitoring, evaluation and review of national health strategies: A country-led platform for information and accountability. Geneva: WHO; 2011.

Additional matter. Health System Responsiveness by N Valentine

World Health Organization, Ministry of Social Affairs and Health of Finland. The Helsinki statement on health in all policies. World Health Organization and Ministry of Social Affairs, Finland; 2013.

World Health Organization, World Bank. Monitoring progress towards universal health coverage: framework, measures and targets. Geneva: World Health Organization; 2014.

World Health Organization, World Bank. Tracking universal health coverage: first global monitoring report. Geneva: WHO; 2015.

World Health Organization. Everybody's business: strengthening health systems to improve health outcomes: WHO's framework for action. Geneva: World Health Organization; 2007.

World Health Organization. Health and human rights and equity working group draft glossary. 2005.

World Health Organization. Health systems: improving performance. The world health report 2000. Geneva: World Health Organization; 2000.

World Health Organization. Measuring and monitoring intersectoral factors influencing equity in universal health coverage (UHC) and health. Summary report of a meeting in Bellagio; 2014 Jul. Report No.: PHE/SDH/2015/1. [online: www.who.int/healthinfo/universal_health_coverage/UHC_Meeting_Bellagio_Mar2014_Report.pdf]

World Health Organization. Reducing risks, promoting healthy life. The world health report 2002. Geneva: World Health Organization; 2002.

World Health Organization. State of Inequality, Geneva: World Health Organization, 2015.

World Health Organization. Technical consultation on concepts and methods for measuring the responsiveness of health systems. In: Murray CJL, Evans DB, editors. Health systems performance assessment debates, methods and empiricism. Geneva: World Health Organization; 2003. p.115–24.

World Health Organization. The health systems responsiveness analytical guidelines for surveys in the Multi-Country Survey Study 2005 [Internet]. Geneva: World Health Organization; 2005 (December) [cited 15 June 2016]. Available from: http://www.who.int/responsiveness/papers/MCSS_Analytical_Guidelines.pdf?ua=1 World Health Organization. World health statistics 2005. Geneva; 2005.

World Health Organization. World health statistics 2005. Geneva; 2005.

World Health Organization. World health statistics 2006. Geneva; 2006.

World Health Organization. World health statistics 2007. Geneva; 2007.

World Health Organization. World health statistics 2008. Geneva; 2008.

World Health Organization: Technical consultation on concepts and methods for measuring the responsiveness of health systems. In: Murray CJL, Evans DB, editors. Health systems performance assessment debates, methods and empiricism. Geneva: World Health Organization; 2003. p.115–24.

World Health Survey [Internet]. Geneva:WHO Global Health Observatory. [cited August 2014]. Available from: <http://www.who.int/healthinfo/survey/en/>

Wouters A. Essential national health research in developing countries: health care financing and the quality of care. *Int J Health Plan M.* 1991;6:253–71.

Zastowny TR, Stratmann WC, Adams EJ, Fox ML. Patient satisfaction and experience with health services and quality of care. *Qual Manag Health.* 1995;C(3):50–61.

Zeitlin J, Wildman K, Breart G, Alexander S, Barros H, Blondel B, Buitendijk S, Gissler M, Macfarlane A. PERI-STAT: indicators for monitoring and evaluating perinatal health in Europe. *Eur J Public Health*. 2003;13(3 Suppl):29–37.

Zimlichman EI, Rozenblum R, Millenson ML. The road to patient experience of care measurement: lessons from the United States. *Isr J Health Pol Res*. 2013;17(2):1.

Zweifler J, Hughes S, Lopez RA. Controlling for race/ethnicity: a comparison of California commercial health plans CAHPS scores to NCBD benchmarks. *Int J Equity Health*. 2010(Jan);9:e4. doi: 10.1186/1475-9276-9-4.

AUTHORS AND AFFILIATIONS



AUTHORS AND AFFILIATIONS

From the Department of Obstetrics and Gynaecology, Division of Obstetrics and Prenatal Medicine, Erasmus MC, University Medical Centre Rotterdam, The Netherlands:

Gouke J. Bonse, Semiha Denктаş, Johanna (Hanneke) P. de Graaf, Eric A.P. Steegers, Jacoba van der Kooy, Marijana Vujkovic

From the Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, and Utrecht University, The Netherlands:

Gouke J. Bonse

From the Institute of Health Policy and Management (current name: Erasmus School of Health Policy & Management), Erasmus University, Rotterdam, The Netherlands:

Erwin Birnie

From the Agency for Healthcare Quality and Research, Baltimore, US:

Charles H. Darby

From the University of Harvard, Cambridge, US:

Christopher J.L. Murray

From the University of York, UK:

Nigel Rice, Silvana Robone

From the World Health Organization, Geneva, Switzerland:

Somnath Chatterji, Amit Prasad, Emese Verdes-Tennant

MANUSCRIPTS



A. MANUSCRIPTS RELATED TO THIS THESIS

SECTION I

Chapter 2

N.B. Valentine, G.J. Bonsel, C.J.L. Murray. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. *Quality of Life Research* 2007, 16(7): 1107 – 25.

Chapter 3

N.B. Valentine, A. Prasad, N. Rice, S. Robone, S. Chatterji. 'Health systems responsiveness: a measure of the acceptability of health-care processes and systems from the user's perspective'. In P.C. Smith, E. Mossialos, I. Papanicolas & S. Leatherman. *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects*. Cambridge University Press: Cambridge, 2009. ISBN 978-0-521-11676-3 Hardback; ISBN 978-0-521-13348-7 Paperback.

Chapter 4

N.B. Valentine, E. Verdes, G.J. Bonsel. Health systems' responsiveness and reporting behaviour: multilevel analysis of individual-level factors contributing to reporting behaviour heterogeneity in 64 countries. *Social Science and Medicine* 2015. Aug;138:152 – 60.

SECTION II

Chapter 5

N.B. Valentine, C. Darby, G.J. Bonsel. Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of "health systems responsiveness" in 41 countries. *Social Science and Medicine* 2008, 66: 1939 – 50.

Chapter 6

N.B. Valentine, G.J. Bonsel. What explains users' reports on health system responsiveness? Exploring health service and personal characteristics across 49 countries. Manuscript submitted for publication.

Chapter 7

N.B. Valentine, G.J. Bonsel. Exploring models for the roles of health systems' responsiveness and social determinants in explaining universal health coverage and health outcomes. *Global Health Action* 2016. [S.I.] 9: 29329 (1 – 16). Available at: <<https://www.tandfonline.com/doi/>

full/10.3402/gha.v9.29329%40zgha20.2016.9.issue-s3>. Date accessed: 26 Mar. 2018.
doi:<http://dx.doi.org/10.3402/gha.v9.29329>.

SECTION III

Chapter 8

J. Van der Kooy J., **N.B. Valentine**, E. Birnie, M. Vujkovic, J. de Graaf, S. Denktas, E.A.P. Steegers, G.J. Bonsel. Validity of a questionnaire measuring the World Health Organization concept of health system responsiveness with respect to perinatal services in the Dutch obstetric care system. *BMC Health Services Research* 2014. 14(622):1 – 13. Available at: <<http://www.biomedcentral.com/1472-6963/14/622>>. Date accessed: 26 March 2018. doi:<http://dx.doi.org/10.1186/s12913-014-0622-1>

Chapter 9

J van der Kooy, E. Birnie, **N.B. Valentine**, J.P. de Graaf, S. Denktas, E.A. P. Steegers, G.J. Bonsel. Quality of perinatal care services from the user's perspective: a Dutch study applies the World Health Organization's responsiveness concept. *BMC Pregnancy and Childbirth*, 17(327): 1 – 11. Available at: < <https://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/s12884-017-1464-8>>. Date accessed: 26 March 2017. doi: <https://doi.org/10.1186/s12884-017-1464-8>

B. OTHER RELATED MANUSCRIPTS ON RESPONSIVENESS

Scientific journals

C.J.L Murray C.J.L., K. Kawabata K., **N.B. Valentine**. People's experience versus people's expectations. *Health Affairs* 2000; 20 (3): 21 – 4.

N.B. Valentine (2011). Contribution to the discussion presented at the Ordinary Meeting of the Royal Statistical Society in Dublin on the paper, 'Vignettes and health systems responsiveness in cross-country comparative analyses' by Nigel Rice and Peter Smith. Published in: *Journal of the Royal Statistical Society Series A (Statistics in Society)* 01/2012; 175(2):357 – 69. doi: <http://dx.doi.org/10.2307/23251886>

Scientific conferences

N.B. Valentine, G.J. Bonsel (2003). Responsiveness for Inpatient and Outpatient Health Services from Population Surveys in 32 Countries: differences by level of GDP and the level of civil society liberty. *International Health Economics Association (IHEA) Conference 2003*, 14-19 June, San Francisco, USA [No web link available; copy of paper available from author; IHEA paper conference programmes for 2003 were destroyed and records of this date were not transferred to digital archives]

J. Poeran, **N.B. Valentine**, G.J. Bonsel, A. Zauber, J. Moore (2014). Statistics in Epidemiology: Communication is the Key to Success. *Joint Statistical Meeting*, 2-7 August 2014. Abstract available at: <<http://www.amstat.org/meetings/jsm/2014/onlineprogram/AbstractDetails.cfm?abstractid=311354>>

Books

N.B. Valentine, A. de Silva, K. Kawabata, C. Darby, C.J.L. Murray, D.B. Evans (2003). "Health Systems Responsiveness: concepts, domains and operationalization". In C.J.L. Murray & D.B. Evans (eds.), *Health Systems Performance Assessment: debates, methods and empiricism*, WHO, Geneva. ISBN: 9241562455

N.B. Valentine, J. Salomon (2003). "Weights for Responsiveness Domains". In C.J.L. Murray & D.B. Evans (eds.), *Health Systems Performance Assessment: debates, methods and empiricism*, WHO, Geneva. ISBN: 9241562455

N.B. Valentine, R. Lavallée, B. Liu, G.J. Bonsel, C.J.L. Murray (2003). "Classical Psychometric Assessment of the Responsiveness Instrument in the WHO Multi-Country Survey Study on Health and Responsiveness 2000–2001". In C.J.L. Murray & D.B. Evans (eds.), *Health Systems Performance Assessment: debates, methods and empiricism*, WHO, Geneva. ISBN: 9241562455

N.B. Valentine, J.P. Ortiz, A. Tandon et al., K. Kawabata, D.B. Evans, C.J.L. Murray (2003). "Patient experiences with health services: population surveys from 16 OECD countries". In C.J.L. Murray & D.B. Evans (eds.), *Health systems performance assessment: debates, methods and empiricism*. Geneva: World Health Organization. ISBN: 9241562455

R. Busse R, **N.B. Valentine**, S. Lessof, A. Prasad, E. van Ginneken. (2012). "Being responsive to citizen's expectations: the role of health services in responsiveness and satisfaction". In J. Figueras & M. McKee (eds.), *Health Systems, Health, Wealth and Societal Well-being: assessing the case for investing in health systems*, Open University Press, Maidenhead. ISBN-13: 978-0-335-24430-0 (pb) ; ISBN-10: 0-335-24430-0; eISBN: 978-0-335-24431-7

Health reports

L. Gostin L, G. Hodge, **N.B. Valentine**, H. Nygren-Krug (2003). The domains of health responsiveness: a human rights analysis. Evidence and information for policy (EIP) discussion paper 53. 2003. [online: http://apps.who.int/iris/bitstream/10665/73926/1/HHRWPS2_eng.pdf]

Web-accessible reports for MCSS. **N.B. Valentine** and A. Prasad.
(Unpublished analysis of World Survey country data) **N.B. Valentine** and A. Prasad. Analytical tables by country for the World Health Survey.

Sample reports.

<http://www.who.int/responsiveness/MCSS%20Sample%20Report.pdf?ua=1>

<http://www.who.int/responsiveness/MCSS%20Data%20Analyst's%20Report.pdf?ua=1>

C. UNRELATED PUBMED PAPERS BY THE AUTHOR

S. Kuruvilla, R Sadana, EV Montesinos, J. Beard, J.F. Vasdeki, I Araujo de Carvalho, R. Thomas, MB Drisse, B. Daelmans, T. Goodman, T Swift Koller, A. Officer, J. Vogel, **N.B. Valentine**, E. Wootton, A. Banerjee, V. Magar, M. Neira, J.M. Bele, A.M. Worning, F. Bustreo. A life-course approach to health: synergy with sustainable development goals. *Bulletin of the World Health Organization* 2018(January); 96(1):42-50. doi: 10.2471/BLT.17.198358. Epub 2017 Nov 23. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5791871/>

F.Pega, **N.B. Valentine**, K. Rasanathan, A.Reza Hosseinpoor, T.P. Torgersen, V. Ramanathan, T. Posayanonda, N. Röbbel, Y. Kalboussi, D.H. Rehkopf, C. Dora, E.R.Villar Montesinos, M.P.Neira. *Bulletin of the World Health Organization* 2017(Nov 1);95(11):784-7. doi: 10.2471/BLT.16.184622. Epub 2017 Sep 12. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5677605/>

N.B. Valentine, T. Swift Koller, A. Reza Hosseinpoor. Monitoring health determinants with an equity focus: a key role in addressing social determinants, universal health coverage, and advancing the 2030 sustainable development agenda. *Global Health Action* 2016; 9. doi: 10.3402/gha.v9.34247. Epub 2016 Dec 16.
<https://www.ncbi.nlm.nih.gov/pubmed/25243463/>

H. Van Minh, K.B. Giang, L.N. Hoat, H. le Chung, T.T. TT5, N.T. Phuong, **N.B. Valentine**. Analysis of selected social determinants of health and their relationships with maternal health service coverage and child mortality in Vietnam. *Global Health Action* 2016 (Feb); 4;9:28836. doi: 10.3402/gha.v9.28836.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4744327/>

A. Reza Hosseinpoor, N. Bergen, T Swift Koller, A. Prasad, **N.B. Valentine**, J. Lynch, J. Vega. Equity-oriented monitoring in the context of universal health coverage. *PLoS Med* 2014; 11(9): e1001727.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4171107/>

F. Pega, **N.B. Valentine**, D. Matheson, K. Rasanathan. Public social monitoring reports and their effect on a policy programme aimed at addressing the social determinants of health to improve health equity in New Zealand. *Social Science and Medicine* 2014; 101:61-9.
<https://www.ncbi.nlm.nih.gov/pubmed/24560225/>

K. Rasanathan, J. Norenhag, **N.B.Valentine**. Realizing human rights-based approaches for action on the social determinants of health. *Health and Human Rights: An International Journal* 2010; 12 (2) : 49-59.
<https://www.ncbi.nlm.nih.gov/pubmed/21178189/>

R. Krech R., **N.B. Valentine**, L. Tucker Reinders, D. Albrecht. Implications of the Adelaide Statement on Health in All Policies. *Bulletin of the World Health Organization* 2010; 88:720-720.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2947051/>

A. Irwin, **N.B. Valentine**, C. Brown, R. Loewenson, O. Solar, et al. The Commission on Social Determinants of Health:Tackling the Social Roots of Health Inequities. *PLoS Medicine* 2006.3(6): e106. doi:10.1371/journal.pmed.0030106.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1459479/>

J.J.Wilkins, P. Folb, **N.B. Valentine**, K.I. Barnes. An economic comparison of chloroquine and sulfadoxine-pyrimethamine as first-line treatment for malaria in South Africa: Development of

a model for estimating direct costs. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2001; 96 (1).

<https://www.ncbi.nlm.nih.gov/pubmed/11926003/>

D. McIntyre, **N.B. Valentine**, J. Cornell. Private Sector Health Care Expenditure in South Africa. *South African Journal of Medicine* 1995(March); 85 (3).

<https://www.ncbi.nlm.nih.gov/pubmed/7777953/>

CURRICULUM VITAE



CURRICULUM VITAE

Nicole started her academic career at the University of Cape Town (UCT) in South Africa. After graduating with a degree in Economics, she worked at UCT in operational research for health economics under Di McIntyre, supporting national, provincial and local government public health initiatives, and teaching. During this period while working she also studied for a Master of Economics degree from the University of Cape Town, and she was awarded a Rotary Scholarship to study abroad. With this award, she completed a Master of Public Health from the University of Washington, Seattle (1996/7).

Nicole joined WHO in December 1999, in order to work on the WHO Health Systems Performance Assessment initiative, from which the work described in this thesis originates. With Amala de Silva she prepared a key informant meeting in December 2000, which was one of many fora where the “responsiveness” concept was discussed in its initial phase over 2000 and 2002. The resulting launch of the World Health Report 2000 was the first time an empirical report of health systems worldwide was created, combining evidence on health (in terms of the by then recent burden-of-disease measures) with evidence on financing and quality of its delivery (service level, responsiveness). She led the responsiveness team for 3-4 years, under the coordination of Christopher Murray and Kei Kawabata, during which time the first wave of household surveys was launched (Multi-Country Survey Study), to be followed soon afterward by second wave of global surveys (the World Health Surveys). During this period she worked with colleagues and key advisors to WHO in the Agency for Healthcare Quality and Research in the USA.

From 2004 to 2008 she served in the Secretariat of the global Commission on Social Determinants of Health (CSDH), under Jeanette Vega and Timothy Evans, where she co-led the work stream dealing with country partners to the CSDH and the organization of several meetings of the Commissioners. For a period of 1 year she was also Acting Coordinator of the Equitable Health Systems and Policy unit in the Health Systems cluster.

Following the CSDH, Her work on social determinants of health increasingly focused on intersectoral policy-making as a natural extension of the Report on the Social Determinants of Health – which advocated working across disciplinary and bureaucratic silos to address the determinants of health, and more recently, on monitoring and evaluation of determinants. Projects under her direction were inter alia: the Adelaide Health in All Policies Conferences (2017, 2010); the Health in All Policies Training Manual (2015); Health Equity through Intersectoral Action – An Analysis of 18 Country Case Studies (2008); a series of Health the SDH Discussion Paper and Sectoral Briefing Series (2010/12). Translation of evidence to practice was the focus

of projects like Making the Economic Case for Addressing the Social Determinants of Health and Supporting Regional Positions on Health in All Policies with an Equity-Focus (for the Helsinki Health Promotion Conference, 2013).

While the initial idea for a PhD emerged in 2002/3 in The Netherlands, personal circumstances, both enjoyable and burdensome, twice interrupted the initial schedule. After a first restart in 2010/11, with the work on the survey with ReproQ at Erasmus, family and health difficulties in 2014-16 again halted progress, but she is pleased to be completing her PhD in 2018.

Nicole is currently a Technical Officer and Acting Coordinator in the Department of Public Health, Environmental and Social Determinants of Health (SDH) in WHO, Geneva. She is leading the WHO's global initiative on training in Health in All Policies, which supports regional training, implementation and standards development and forms part of implementation of WHO's Country Framework for Action Across Sectors. She is also coordinating the WHO initiative to develop a monitoring framework and report for action on the social determinants of health.

PhD PORTFOLIO



PhD PORTFOLIO

Summary of PhD related training and teaching activities

| | |
|------------------------|--|
| Name of PhD candidate: | Nicole B. Valentine |
| Erasmus MC Department: | Public Health |
| PhD period: | 2003-2012 (Amsterdam MC); 2013-2018 (transferred to Erasmus MC in 2012/3) |
| Promotor: | Prof dr G. J. Bonsel |

| 1. PhD TRAINING | Years | Workload | |
|---|-------|-------------|------|
| | | Hours | ECTS |
| <u>Specific courses (following Masters in Public Health, Masters in Economics)</u> | | | |
| Equity gauge training | 2004 | 24 hours | |
| Equity in access analysis workshop (Erasmus and University of Lausanne, E van Doorslaer) | 2004 | 8 hours | |
| Joint Statistical Meeting Professional Development program, Meta-Analysis: combining the results of multiple studies | 2014 | 8 hours | |
| <u>General academic skills</u> | | | |
| WHO short courses on scientific writing | 2003 | 21 hours | |
| Grant proposal writing workshop | 2008 | 14 hours | |
| WHO course on database searches | 2010 | 7 hours | |
| <u>Presentations</u> | | | |
| <u>International conferences</u> | | | |
| International Health Economics Association, San Francisco, USA: -oral presentation, parallel session: <i>Responsiveness for Inpatient and Outpatient Health Services from Population Surveys in 32 Countries: differences by GDP and the level of civil society liberty</i> ; -oral presentation, 2 workshop sessions: <i>Levels of responsiveness in eight domains for outpatient and inpatient experiences: developing a cross-population comparable measure</i> ; <i>WHO on Health System Responsiveness and an Introduction to Inequality</i> | 2003 | | 5 |
| Attending the Third International Conference of the International Society for Equity in Health, Durban, South Africa | 2004 | | 0.5 |
| International Union for Health Promotion and Education, Vancouver, Canada: -oral presentation <i>Intersectoral action: Innovative approaches for health and equity</i> -co-author of accepted oral presentation <i>Monitoring social determinants of health</i> | 2007 | | 1 |
| Conference on Quality in Health Care in low and middle income countries, Manila, Philippines: oral presentation <i>Global overview of health system responsiveness: Building equitable, people-centred health systems: what can be learnt about quality of care from WHO's responsiveness work?</i> | 2007 | | 1 |

| 1. PhD TRAINING continued | Years | Workload | |
|--|-------|----------|------|
| | | Hours | ECTS |
| Third EQUINET Regional Conference on Equity in Health in East and Southern Africa - <i>Equity in health and health systems</i> , Kampala, Uganda: attended and plenary welcome | 2009 | | 1 |
| International Union for Health Promotion and Education, Geneva, Switzerland: organized and gave oral presentations for workshop on Building Global Capacity for Health in All Policies and Intersectoral Action | 2010 | | 1 |
| International Conference on Health Impact Assessment, Granada, Spain: oral presentation in plenary <i>The Conceptual Framework for Social Determinants of Health: which theory is the basis for a tool for Health Impact Assessment?</i> | 2011 | | 1 |
| International Union for Health Promotion and Education Conference, October, Pattaya, Thailand: oral presentation in panel <i>Evaluation in Health in all Policies</i> | 2013 | | 1 |
| Joint Statistical Meeting, Boston, USA: oral presentation in panel <i>Reporting health systems responsiveness and patient-centred care: lessons for making qualitative information a spur for action</i> | 2014 | | 1 |
| Global Symposium on Health Systems Research, Cape Town, South Africa: oral presentation <i>How does responsiveness reporting behaviour differ for different social groups and what does this mean for equity measures? Results from cross-sectional household survey data from 64 countries</i> | 2014 | | 1 |
| International meetings | | | |
| Scientific Research Institute of Social Hygiene, Russian Academy of Medical Sciences (SRISH RAMS): 2nd International Seminar on Patient's Rights "The role of civil society and societal structures in promoting Patient's Rights": oral presentation <i>Timely topics in patient rights: human rights and health systems responding to people's needs with caring, dignified care (responsiveness)</i> | 2003 | | 1 |
| Eschborn (GTZ) Dialogue on good governance and the workshop on making social services work for the poor, Berlin, Germany: oral presentation in panel <i>Responsiveness and equity</i> | 2004 | | 1 |
| Eastern Mediterranean Regional Consultation Following up on Mexico Ministerial Summit on Health Research oral presentation <i>Research to promote action on health inequities and poverty</i> | 2005 | | 1 |
| Attended Equity in Access in ART and maternal health programmes: National meeting in Zimbabwe with Ministries of Health for maternal health and AIDS programmes | 2006 | | 1 |
| UNICEF Consultation on Equity: Increasing the Contribution of Health Systems to health equity, New York, USA: oral presentation <i>Addressing health equity through the redistributive role of health systems: learnings from WHO and its Commission on SDH</i> | 2008 | | 1 |
| OECD Health Care Quality Indicators Subgroup on Measurement of Patient Experience/ Responsiveness: teleconference oral presentation <i>Responsiveness/patient satisfaction, and the use of vignettes</i> | 2009 | | 1 |
| The Second Inter-Ministerial Conference on Health and Environment in Africa, Lukada, Angola: oral presentation on <i>Accelerating Response to the Determinants of Health in the African Region</i> | 2010 | | 1 |
| Ordinary Meeting of the 58th World Statistics Congress of the International Statistical Institute, Dublin, Republic of Ireland oral comment on the paper, "Vignettes and health systems responsiveness in cross-country comparative analyses" by Nigel Rice and Peter Smith. | 2011 | | 1 |

Additional matter. Health System Responsiveness by N Valentine

| 2.TEACHING ACTIVITIES | Years | Workload | |
|--|-----------|------------------|-------------|
| | | Hours | ECTS |
| Teaching in workshops | | | |
| WHO Responsiveness data analysis training workshop for 15 countries | 2003 | 56 hours | |
| International Health Programme of the University of Geneva, Geneva, Switzerland: lecture on <i>Policy-maker perspective on working across sectors to improve health and health equity</i> | 2015 | 5 hours | |
| Writing of national responsiveness reports with other staff | | | |
| 61 country survey reports (Multi-Country Survey Study) and production of cleaned datasets and production of cleaned datasets | 2004-5 | 56 hours | |
| Country tables for 68 World Health Surveys | 2005-6 | 56 hours | |
| Supervisory / Advisory work | | | |
| PhD Thesis review- George Shakarsvili: University of Oxford, UK: <i>Analysing the Equity of Post-Soviet Health Care Systems: Evaluation of 1990s Health Reforms</i> | 2004-5 | 7 hours | |
| OECD Health Care Quality Indicators Subgroup on Measurement of Patient Experience/Responsiveness | 2009-2012 | 14 hours | |
| Berlin University of Technology European Union funded project, <i>Exploring health systems responsiveness in ambulatory care and disease management and its relation to other dimensions of health systems performance</i> | 2012-2014 | 14 hours | |
| Master's Thesis - Marie-Cristina Dankmeyer, Master of Public Policy, UCLA Luskin School of Public Affairs: Health in All Policies | 2012-13 | 28 hours | |
| TOTAL | | 318 hours | 21.5 |

WORD OF THANKS



WORD OF THANKS

The first acknowledgement of gratitude goes to my supervisor and promoter, Gouke Bonsel, for staying the course with his student through the various changes in our respective careers. I am also grateful to his wife, Karin, for her hospitality in the times I needed to stay in Holland. A huge debt is owed to my husband, Grant, and our two children, for their patience with my “hobby”, which took me away from household chores and from time otherwise spent together. I would also like to thank my parents, whose combined love nurtured and supported my academic leanings, as well as the Wroe-Streets, my sister, aunts and uncles, cousins and other relatives, who are all so important in my life. They gave me strength, especially Steve and Evelyn. These fine people, good neighbours (old and new) and many dear friends, in particular Keeley, Marie-Louise, Tracey, Orielle, Moira, Susan, Elske, the Woods, Heidi, Thalassa, Tanya, and many others I do not name, were a great support through the years, encouraging me to complete the work.

Various colleagues at the World Health Organization were particularly supportive of my efforts to complete a doctoral thesis while working. I wish to acknowledge with thanks the following colleagues and supervisors who encouraged me over the years. In chronological order, thanks go to: Christopher Murray, Kei Kawabata, Amala De Silva, Charles Darby (RIP) and Susan Darby, Ajay Tandon, Jeanette Vega, David Evans, Timothy Evans, Somnath Chatterji, Amit Prasad, Rüdiger Krech, Lucy Mshana, Nathalie Roebbel, Eugenio Villar and Maria Neira. Thanks go to all my co-authors as well for working on the various papers with me.

Work turned to fun in the last few years with the collegial qualities of Babs van der Kooy and Jashvant Poeren. Also remembered with affection is the kindness and support received in finalizing the manuscript and with organizing the Defence ceremony - to Erwin Birnie, Bas Janssen, Sanne Hanneke de Graf, Sanne Ruseker, and my paranymphs, Nynke de Groot and Moira Potgieter.

Thanks also go to the thesis examination committee members for taking the trouble to read this lengthy manuscript: in the small committee: Prof. dr. E.W. Steyerberg, Erasmus MC; Prof. dr. E.K.A. van Doorslaer, Erasmus School of Economics, Applied Economics and Erasmus School of Health Policy and Management; Prof. dr. A. Franx, UMC Utrecht; and in the large committee: Prof. Dr. D.M.J. Delnoij, Erasmus School of Health Policy & Management; dr. E. Birnie, University of Groningen; Prof. P.C. Smith, Imperial College London. Peter’s wisdom and generosity of spirit over the decades of work on health system performance assessment, of which health system responsiveness was part, have been highly appreciated by many.

Thank you all for some splendid moments and this meaningful adventure.



ANNEX

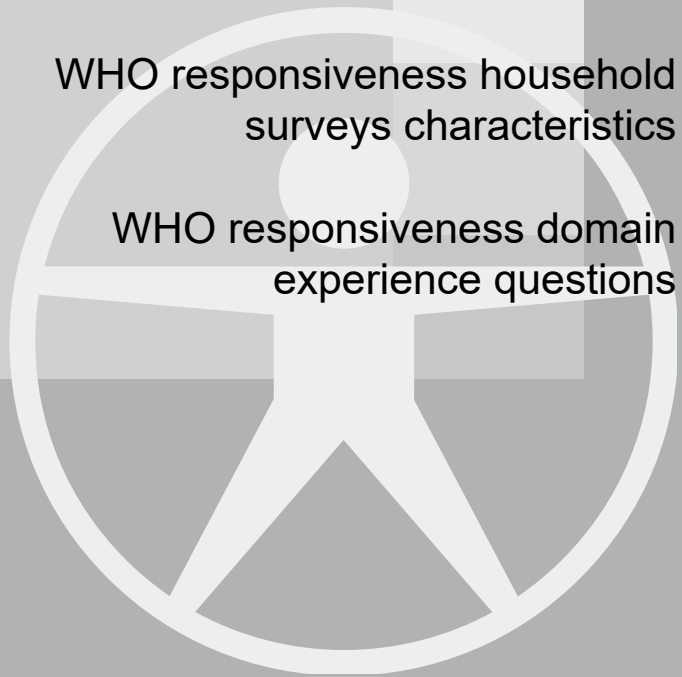
ANNEX A

WHO responsiveness survey development

Annex A1. WHO responsiveness survey countries in this thesis

Annex A2. WHO responsiveness household surveys characteristics

Annex A3. WHO responsiveness domain experience questions



ANNEX A1. WHO RESPONSIVENESS SURVEY COUNTRIES IN THIS THESIS

1. (1999): first household (3 country) test survey: 3: Colombia, Philippines, Tanzania.
2. (2000): pilot multi-country household survey on health and health system responsiveness: 8: China, Colombia , Egypt , Georgia , India , Nigeria, Slovakia , Turkey.
3. (2000-01): the Multi-Country Survey Study on Health and Health Systems Responsiveness: 70 surveys with responsiveness, 41 in-person administered; 28 postal or drop-off.

| Extended Form: (face-to-face) 90 minute questionnaire Contractors: INDEPENDENTS | Short Form: (face-to-face, except Luxembourg.), 60 minute questionnaire, Contractors: INRA | Short Form: (face-to-face) 30 minute questionnaire Contractors: GALLUP | Short Form: (telephone and face-to-face) Contractors: INDEPENDENTS | Short Form: (postal and drop-off) 30 minute questionnaire Contractors: INDEPENDENTS | |
|---|---|---|--|--|---|
| China† Colombia Egypt Georgia India† Indonesia Iran Lebanon Mexico Nigeria† Singapore+ Slovakia Syria Turkey | Belgium Bulgaria Czech Republic Estonia Finland France Germany Iceland Ireland Italy Luxembourg (telephone) Malta Netherlands Portugal Romania Russian Federation Spain Sweden | Argentina Bahrain Costa Rica Jordan Latvia Morocco Oman UAE Venezuela | Canada (30 minute) Croatia (60 minute) | Australia Austria Canada Chile <i>China*†</i> Cyprus Czech Republic Denmark <i>Egypt*</i> Finland France Greece Hungary Indonesia Kyrgyzstan | Lebanon Lithuania Netherlands New Zealand Poland Rep. of Korea Switzerland Thailand Trinidad and Tobago <i>Turkey*</i> Ukraine United Kingdom USA |
| 13 (14, incl. Singapore) | 18 | 9 | 2 | 28 | |

† Not full national sample; +Not collected information on Responsiveness module; *Drop-and-questionnaire

ANNEX A1. WHO RESPONSIVENESS SURVEY COUNTRIES IN THIS THESIS (CONTINUED, 2)

4. (2002-04) the World Health Survey: 70 surveys with responsiveness; 68 surveys' responsiveness questionnaires reported to WHO with 65 surveys covered in thesis: 16 Low income countries; 12 lower middle income countries; 15 higher middle-income countries; 22 high income countries.

64 surveys qualifying for largest analysis of vignettes: Brazil, Hungary, and Zimbabwe omitted the question on an individual-level covariate used in the final model - the importance of responsiveness relative to other health system goals; Belgium contained only 7 respondents in one category of the "intensity/ type of experience" (individual-level covariate used in the Chapter 7 vignette regression model).

| AFRICA | AMERICAS | EUROPE | EASTERN MEDITERRANEAN | SOUTH-EAST ASIA | WESTERN PACIFIC |
|---|--|---|--|---|---|
| Burkina Faso Chad Comoros† Congo† Côte d'Ivoire† Ethiopia Ghana Kenya Malawi Mali Mauritania Mauritius Namibia Senegal South Africa Swaziland Zambia Zimbabwe | Brazil Chile+ Dominican Rep., Ecuador Guatemala Mexico Paraguay Uruguay | [Austria Belgium Denmark Finland France Germany Greece Ireland Italy Luxembourg* Netherlands Portugal Sweden United Kingdom] Bosnia and Herzegovina Croatia Czech Rep. Estonia Georgia Hungary Israel* Kazakhstan Latvia Norway* Russian Fed.† Slovakia Slovenia Spain Ukraine | Morocco Pakistan Tunisia United Arab Emirates | Bangladesh India† Myanmar Nepal Sri Lanka | Australia+* China† Lao (PDR) Malaysia Philippines Viet Nam |
| 18 | 8 | 29 | 4 | 5 | 6 |

† Not full national sample; *Computer-assisted telephone interviews; [Used a common survey platform with the short questionnaire]; +Not reporting Responsiveness module to WHO

ANNEX A1. WHO RESPONSIVENESS SURVEY COUNTRIES COVERED IN THIS THESIS (CONTINUED, 3)

5. Combined set of MCS Study and WHS country surveys analyzed (maximum)

| Number | Survey country | MCS | WHS | Both |
|--------|------------------------|-----|-----|------|
| 1 | Argentina | x | | |
| 2 | Austria | | x | |
| 3 | Bahrain | x | | |
| 4 | Bangladesh | | x | |
| 5 | Belgium | x | x | X |
| 6 | Bosnia and Herzegovina | | x | |
| 7 | Bulgaria | x | | |
| 8 | Burkina Faso | | x | |
| 9 | Canada | x | | |
| 10 | China | x | x | x |
| 11 | Colombia | x | | |
| 12 | Comoros | | x | |
| 13 | Congo | | x | |
| 14 | Costa Rica | x | | |
| 15 | Cote d'Ivoire | | x | |
| 16 | Croatia | x | x | x |
| 17 | Czech Rep. | x | x | x |
| 18 | Denmark | | x | |
| 19 | Dominican Republic | | x | |
| 20 | Ecuador | | x | |
| 21 | Egypt | x | | |
| 22 | Estonia | x | x | x |
| 23 | Ethiopia | | x | |
| 24 | Finland | x | x | x |
| 25 | France | x | x | x |
| 26 | Georgia | x | x | x |
| 27 | Germany | x | x | x |
| 28 | Ghana | | x | |
| 29 | Greece | | x | |
| 30 | Guatemala | | x | |
| 31 | Iceland | x | | |
| 32 | India | x | x | x |
| 33 | Indonesia | x | | |
| 34 | Iran | x | | |

ANNEX A1. WHO RESPONSIVENESS SURVEY COUNTRIES COVERED IN THIS THESIS (CONTINUED, 4)

5. Combined set of MCS Study and WHS country surveys analyzed (maximum) (continued, 2)

| Number | Survey country | MCS | WHS | Both |
|--------|-------------------|-----|-----|------|
| 35 | Ireland | x | x | x |
| 36 | Israel | | x | |
| 37 | Italy | x | x | x |
| 38 | Jordan | x | | |
| 39 | Kazakhstan | | x | |
| 40 | Kenya | | x | |
| 41 | Laos | | x | |
| 42 | Latvia | x | x | x |
| 43 | Luxemburg | x | x | x |
| 44 | Malawi | | x | |
| 45 | Malaysia | | x | |
| 46 | Mali | | x | |
| 47 | Malta | x | | |
| 48 | Mauritania | | x | |
| 49 | Mauritius | | x | |
| 50 | Mexico | x | x | x |
| 51 | Morocco | x | x | x |
| 52 | Myanmar | | x | |
| 53 | Namibia | | x | |
| 54 | Nepal | | x | |
| 55 | Netherlands (The) | x | x | x |
| 56 | Nigeria | x | | |
| 57 | Norway | | x | |
| 58 | Oman | x | | |
| 59 | Pakistan | | x | |
| 60 | Paraguay | | x | |
| 61 | Philippines | | x | |
| 62 | Portugal | x | x | x |
| 63 | Romania | x | | |
| 64 | Russia | x | x | x |
| 65 | Senegal | | x | |
| 66 | Slovakia | x | x | x |
| 67 | Slovenia | | x | |
| 68 | South Africa | | x | |

ANNEX A1. WHO RESPONSIVENESS SURVEY COUNTRIES COVERED IN THIS THESIS (CONTINUED, 5)

5. Combined set of MCS Study and WHS country surveys analyzed (maximum) (continued, 3)

| Number | Survey country | MCS | WHS | Both |
|--------|----------------------|-----------|-----------|-----------|
| 69 | Spain | x | x | x |
| 70 | Sri Lanka | | x | |
| 71 | Swaziland | | x | |
| 72 | Sweden | x | x | x |
| 73 | Syria | x | | |
| 74 | Trinidad and Tobago | | x | |
| 75 | Tunisia | | x | |
| 76 | Turkey | x | | |
| 77 | UK (England) | | x | |
| 78 | Ukraine | | x | |
| 79 | United Arab Emirates | x | x | x |
| 80 | Uruguay | | x | |
| 81 | Viet Nam | | x | |
| 82 | Zambia | | x | |
| 83 | Venezuela | x | | |
| | Total | 41 | 65 | 23 |

ANNEX A2. WHO RESPONSIVENESS HOUSEHOLD SURVEYS CHARACTERISTICS

| Surveys | 1 (2000): Pilot household surveys | 2 (2000-2001): WHO Multi-Country Survey Study on Health and Health System Responsiveness | 3 (2002-2004): World Health Surveys |
|---|---|---|---|
| Module Characteristics | | | |
| No. of surveys | 8 countries /8 surveys. | 60 countries (70 surveys in total including responsiveness; 71 surveys in total but Singapore excluded the responsiveness module). | 70 countries (70 surveys in total including responsiveness – from original 71 in total: Turkey did not include responsiveness [although planned as 72, Yugoslavia was not pursued]; 68 surveys reported on responsiveness to WHO: the 2 not reporting to WHO: Australia (module was altered), Chile (not reported). |
| No. of surveys for thesis analysis (maximum) | 6 surveys (analysed for importance results to compare ranking (<i>chapter 5</i>)). | 41 surveys with interviewer administered questionnaires. 28 surveys were postal/drop-collect. 1 country, survey was not retrieved (Lebanon). | 65 surveys covered across different analyses: largest analysis (vignettes, <i>chapter 4</i>) had 64 surveys with interviewer administered questionnaires: (excluded 4 of 68 surveys with responsiveness were excluded in largest analysis on vignettes: they omitted the question on the importance of responsiveness relative to other health system goals (Brazil, Hungary, Zimbabwe); or for reasons of too few observations in one cell (threshold of 15 for each cell): Belgium and was excluded). |
| No. of respondents for thesis analysis (maximum) | 59-152 per survey totaling 811. | 348- 9,952 per survey; Records analysed: 105,806 (41 interviewer administered surveys) | 600-10,000 per survey; Records analysed: 152,445 (65 interviewer administered surveys covered in total in thesis including analysis of Belgium, n=597 (<i>chapter 3</i>); 151,848; (64 interviewer administered surveys for largest analysis (<i>chapter 4</i>)). |
| Survey mode and sampling | Face-to-face household survey; Kish tables used for selection of household respondent in some cases; in other cases, whoever was home and would answer the questionnaire. Convenience sample of respondents from sampling frame to be used for main survey. Sites purposively selected respondents to get an even distribution across different population sub-groups: urban/ rural, sex, high and low education, age groups. | Face-to-face, telephone, and postal/drop-and-collect household surveys; extended and brief versions of the module; used in long and brief questionnaires. Kish tables or "last birthday" methods most commonly used for selection of respondents from within households. Sampling designs: generally stratified, multi-stage random sampling for face-to-face surveys – see details below. A 90-minute long version of the interview and a shorter 30-minute version were used. | Mostly face-to-face with long and short questionnaires. Of 57 countries with sampling weights, 51 were nationally representative. They were geographically representative in China, Comoros, Congo, Côte d'Ivoire, India, and the Russian Federation and Guatemala did not report survey sampling weight data. Australia, Israel, Luxembourg, and Norway implemented the WHS short form through CATI (Computerized Telephone Interviews). All modules could be fielded in an average of 90 minutes and a 30-minute version was developed for countries where costs of a 90-minute interview would be prohibitive. |

ANNEX A2. WHO RESPONSIVENESS HOUSEHOLD SURVEYS CHARACTERISTICS (CONTINUED, 1)

| Surveys | 1 (2000): Pilot household surveys | 2 (2000-2001): WHO Multi-Country Survey Study on Health and Health System Responsiveness | 3 (2002-2004): World Health Surveys |
|--|--|--|--|
| Module Characteristics | | | |
| Ethical clearance | Study protocols and processes were cleared by the WHO Sub-Committee for Research Involving Human Subjects and respondent consent was obtained before interviewing. | Study protocols and processes were cleared by the WHO Sub-Committee for Research Involving Human Subjects and respondent consent was obtained before interviewing. | Initial ethical approval was obtained from an independent ethics review conducted by the Harvard School of Public Health's Institutional Review Board. Ethical clearance was obtained from the relevant ethics committee at each site. Respondent informed consent was obtained before interviewing. |
| Translation and pre-testing | Translated into 1 official language per country according to the WHO Translation Guidelines. Translations and back-translations of key terms checked by WHO. Cognitive interviews (delayed retrospective probing) were conducted in 7 sites (China, Egypt, Georgia, India, Nigeria, Slovakia.- poor completion in 1 site (Indonesia, 5)). Each of 6 sites interviewed 20-33 respondents. After the pilot, unclear questions and translation problems were identified and more suitable terms were substituted. | Revised questions were translated into at least 1 official language per country according to the WHO Translation Guidelines. Translations and back-translations of key terms checked by WHO. | A standardized protocol including back translation was used. Quality was independently verified by bilingual experts. Between February and April 2002, pilots including the responsiveness module were conducted using non-random sampling (total respondents= 3639): Cote d'Ivoire (598), India (649), Malaysia (602), Mexico (604), South Africa (585), Spain (592). Having already tested most domain questions in the MCS Study, the pilot phase, was used more to structure the final questionnaire, i.a. to focus on the last event rather than multiple events. |
| Test-retests | Yes, no further information available. | Test-retests for 9 countries: 2174 (58-412) ambulatory care interviews; 183 (0-56) home care interviews; 283 (6-64) inpatient interviews. | 1,200 respondents from the piloted surveys of the 3,630 respondents were retested within two weeks. |
| Q by Q | Yes | Yes | Yes |
| Domains of patient experience (labels in questionnaire order – for item handles, see below) | 8: dignity, autonomy, confidentiality, communication, prompt attention, access to social support, quality basic amenities, choice. | 8: dignity, autonomy, confidentiality, communication, prompt attention, access to social support, quality basic amenities, choice. | 8: dignity, autonomy, confidentiality, communication, prompt attention, access to social support, quality basic amenities, choice. |

ANNEX A2. WHO RESPONSIVENESS HOUSEHOLD SURVEYS CHARACTERISTICS (CONTINUED, 2)

| Surveys | 1 (2000): Pilot household surveys | 2 (2000-2001): WHO Multi-Country Survey Study on Health and Health System Responsiveness | 3 (2002-2004): World Health Surveys |
|---|---|--|---|
| Module Characteristics | | | |
| <u>Responsiveness experience descriptions:</u> number of questions and response options | Questions with responses measured in units of time: 4 (ambulatory), 1 (home care), 1 (inpatient); questions with "never" to "always" responses: 12 (ambulatory), 9 (home care) and 2 (whole health system); rating from 0-10: 8 (ambulatory), 7 (home care), 3 inpatient, 3 whole health system; "yes"- "no" questions 2 (ambulatory), 1 (inpatient); "not a problem", "somewhat of a problem", "quite a problem" responses: 2 (ambulatory), 1 (home care), 2 (hospital); "very poor", "poor", "good", "very good": 2 (ambulatory), 3 (inpatient) | Questions with responses measured in units of time: 2 (ambulatory care – 1 categorical, 1 continuous) 2 (home care*); questions with "never" to "always" responses: 11 (ambulatory), 13 (home*); questions with "no problem", "mild problem", "moderate problem", "severe problem", "extreme problem": 2 (ambulatory), 2 (home*), 2 (inpatient); questions with "very bad", "bad", "moderate", "good", "very good": 9 (ambulatory), 6 (home*), 8 (inpatient); "yes"- "no" questions: 1 (inpatient) <i>*home care questions in long version (13 countries)</i> | Questions with responses measured in units of time: 2 (1 ambulatory care - continuous, 1 inpatient – continuous); questions with responses as number of beds 1 (inpatient – integer); questions with "very good" to "very bad" responses: 13 (ambulatory) 15 (inpatient); |
| <u>Responsiveness valuations / preferences (importance):</u> question and response options | Ranking of all 8 domains from most important to least important (ties permitted) | Asked to say which is the most and least important domain (ties permitted). | Asked to rate the domain on a verbal response scale: Extremely Important; Very Important; Moderately Important; Slightly Important; Not Important |
| <u>Responsiveness vignettes:</u> number of questions and response options | None | 7 questions per domain; 2 domains per sets of 4 rotated across the sample (i.e. 25 percent of sample responded to each set). Response options: "very bad" to "very good" | 10 questions per domain; 2 domains per sets of 4 rotated across the sample (i.e. 25 percent of sample responded to each set). Response options: "very bad" to "very good" |

ANNEX A2. WHO RESPONSIVENESS HOUSEHOLD SURVEYS CHARACTERISTICS (CONTINUED, 3)

| Surveys | 1 (2000): Pilot household surveys | 2 (2000-2001): WHO Multi-Country Survey Study on Health and Health System Responsiveness | 3 (2002-2004): World Health Surveys |
|---|---|---|---|
| Module Characteristics | | | |
| Module Introduction | <p><u>Whole module:</u> "Now I would like to ask you some questions about where you go for health care. First, I will ask you about places you go for health care, where you do not stay overnight to receive care. I will also ask you about the doctors or other health care providers you see there. I will also ask you about health care you receive in your home." <u>Ambulatory:</u> Please tell me the name of the place or person you visit <u>most often</u> for health care. This may be a clinic, hospital or a person you go to for care. The person may be a medical doctor, nurse, pharmacist or person who practices traditional medicine. We need this information to follow up with the health care provider to find out more about their facility and services. You will not be identified to the provider in any way. <u>Inpatient:</u> "Now I would like to ask you some questions about getting health care from a place where you stay over night, which in most cases are hospitals."</p> | <p><u>Whole module:</u> These questions are about your experiences in getting health care in the last 12 months. "This may be from a doctor's consulting room, a clinic, a hospital or a health care provider may have visited you at home." <u>Ambulatory:</u> no specific introduction, flows from questions about last visit to ambulatory care setting <u>Home care:</u> Now for all the following questions on health care you receive at home, I would like you to think about all the health care providers who visited you at home over the last <u>12 months</u>. <u>Inpatient:</u> Now I would like to ask you some questions about getting health care from a place where you stay over-night, which in most cases are hospitals.</p> | <p><u>Whole module:</u> These questions are about when there was last a need for health care for the respondent or for a child of 12 years or less and whether health care was received. <u>Ambulatory:</u> asked whether they used health care services in the last 12 months for themselves or a child (defined as above): "What was the name of the last health care provider you [your child] used in the last 12 months?" [Interviewer: try get the name of the clinic or health centre, rather than the doctor, if the respondent used a clinic or health centre. If the respondent was visited at home, write "home visit"]. This may be from a doctor's consulting room, a clinic, a hospital or a health care provider may have visited you at home. Then asked to categorize the type of healthcare provider from a closed list. <u>Inpatient:</u> Using the same framing as above for a health experience for the respondent and or a child, "What was the name of the last hospital or long term care facility you [your child] stayed in, in the last 5 years?"; "When was your [child's] last overnight stay?" [Interviewer: stop reading further as soon as the respondent has selected one] 1. In the last 4 weeks; 2. In the last year; 3. In the last 2 years; 4. In the last 3 years; 5. In the last 5 years.</p> |
| Type of health services covered: ambulatory, inpatient, home care, other; (recall) | Ambulatory care, home care, inpatient care, and the "whole health system" – all dealt with separately (recall: ambulatory care: 6 months; inpatient: 12 months; whole health system: 12 months). | Ambulatory care, home care, inpatient care, and the "whole health system" (discrimination and financial barriers to care) (recall: 12 months). | Ambulatory care, inpatient care, and the "whole health system" (discrimination and financial barriers to care) (recall: 12 months outpatient; up to 5 years in patient; cut-off 3 years used for analyses). |

ANNEX A3. WHO RESPONSIVENESS DOMAIN EXPERIENCE QUESTIONS

| Responsiveness Domains | | World Health Survey * | | Multi-Country Survey Study |
|---------------------------|--------------------------------|---|--|---|
| Domain Label (alphabetic) | Short Description | Items For Patients and Close Others (as parents) | Pilot Items for World Health Survey | Items For Patients (order follows questionnaire) |
| Autonomy | Involvement in decisions | | | how often did doctors, nurses or other health care providers involve you in deciding about the care, treatment or tests |
| | | how would you rate your experience of getting information about other types of treatments or tests | did you encounter any problem getting involved as much as you wanted to be in making decisions about your health care or treatment | how often did doctors, nurses or other health care providers ask your permission before starting the treatment or tests |
| | | *how would you rate your experience of being involved in making decisions about your health care or treatment | did you encounter any problem getting information about other types of tests or treatment as much as you wanted | *rate your experience of getting involved in making decisions about your care or treatment |
| Choice | Choice of health care provider | | how big a problem was it to see a health care provider you were comfortable with | health care providers available to you how big a problem, if any, was it to get a health care provider you were happy with |
| | | | how big a problem was it to seek a second opinion of your medical condition if you wanted | how big a problem, if any, was it to get to use other health services other than the one you usually went to |
| | | *how would you rate the freedom you had to choose the health care providers that attended to you | | *how would you rate your experience of being able to use a health care provider or service of your choice |
| Communication | Clarity of Communication | | | how often did doctors, nurses or other health care providers listen carefully to you |
| | | *how would you rate the experience of how clearly health care providers explained things to you | did you experience any problem with understanding the way doctors, nurses or other health care providers explained things | *how often did doctors, nurses or other health care providers, explain things in a way you could understand |
| | | *how would you rate your experience of getting enough time to ask questions about your health problem or treatment | did you experience any problem with getting enough time to ask questions about your health problem or treatment | *how often did doctors, nurses, or other health care providers give you time to ask questions about your health problem or treatment |
| | | | | rate your experience of how well health care providers communicated with you in the last 12 months |

ANNEX A3. WHO RESPONSIVENESS DOMAIN EXPERIENCE QUESTIONS (CONTINUED, 1)

| Responsiveness Domains | | World Health Survey * | | Multi-Country Survey Study |
|----------------------------|---|--|---|--|
| Domain Label (alphabetic) | Short Description | Items For Patients and Close Others (as parents) | Pilot Items for World Health Survey | Items For Patients (order follows questionnaire) |
| Confidentiality | Confidentiality of personal information | *how would you rate the way the health services ensured you could talk privately to health care providers | did you experience any problem with being able to talk privately to your doctor, nurse or other health care provider so other people could not overhear your conversation | *how often were talks with your doctor, nurse or other health care provider done privately so other people who you did not want to hear could not overhear what was said |
| | | *how would you rate the way your personal information was kept confidential | did you encounter any problem with having your medical history or reason for your visit kept confidential | *how often did your doctor, nurse or other health care provider keep your personal information confidential? This means that anyone whom you did not want informed could not find out about your medical conditions |
| Dignity | Respectful treatment and communication | *how would you rate your experience of being greeted and talked to respectfully | was it a problem for you to be greeted and talked to respectfully by doctors, nurses or other health care providers | *how often did doctors, nurses or other health care providers treat you with respect |
| | | | | how often did the office staff, such as receptionists or clerks there, treat you with respect |
| | | *how would you rate the way your privacy was respected during physical examinations and treatments | was there any problem with the way your privacy was respected during your physical examinations and treatments | *how often were your physical examinations and treatments done in a way that your privacy was respected |
| | | | | rate your experience of being treated with dignity |
| Quality of basic amenities | Physical surroundings | *how would you rate the cleanliness of the rooms inside the facility, including toilets | how would you rate the cleanliness of the surroundings (including toilets if you used them) | *how would you rate the cleanliness of the place |
| | | *how would you rate the amount of space you had | how would you rate the conditions, such as space and ventilation, in the waiting and examination rooms | *how would you rate the basic quality of the waiting room, for example, space, seating and fresh air |
| | | | | rate the quality of the surroundings, for example, space, seating, fresh air and cleanliness of the health services |

ANNEX A3. WHO RESPONSIVENESS DOMAIN EXPERIENCE QUESTIONS (CONTINUED, 2)

| Responsiveness Domains | | World Health Survey * | | Multi-Country Survey Study |
|--|--|---|--|--|
| Domain Label (alphabetic) | Short Description | Items For Patients and Close Others (as parents) | Pilot Items for World Health Survey | Items For Patients (order follows questionnaire) |
| Prompt attention | Convenient travel and short waiting times | how would you rate the travelling time | how much of a problem was it to travel to a healthcare provider | how often did you get care as soon as you wanted |
| | | *how would you rate the amount of time you waited before being attended to | how much of a problem was it to get attended to by the health care provider quickly | *how would you rate your experience of getting prompt attention at the health services |
| Access to family and community support | Contact with outside world and maintenance of regular activities | *how would you rate the ease of having family and friends visit you | how much of a problem was it for you [your child] to have family and friends visit you | *how big a problem, if any, was it to get the hospital to allow your family and friends to take care of your personal needs, such as bringing you your favourite food, soap etc.. |
| | | how would you rate your [child's] experience of staying in contact with the outside world when you [your child] were in hospital? | how much of a problem was it to keep in contact with the outside world | how big a problem, if any, was it to have the hospital allow you to practice religious or traditional observances if you wanted to |
| | | | | how would you rate your experience of how the hospital allowed you to interact with family, friends and to continue your social and/ or religious customs |

* Similar items in World Health Survey and Multi-Country Survey Study

ANNEX B

WHO Responsiveness questionnaires

Annex B1. Multi-Country Survey Study
Questionnaire on Responsiveness
<http://www.who.int/responsiveness/surveys/en/>

Annex B2. World Health Survey Questionnaire
on Responsiveness
<http://www.who.int/healthinfo/survey/instruments/en/>



Annex B1. Multi-Country Survey Study Questionnaire on Responsiveness

F. Health System Responsiveness

Read all options to the respondent except for Refuse and Don't Know (DK). If a question does not apply to the respondent, circle the option Not Applicable (NA).

These questions are about your experiences in getting health care in the last 12 months. This may be from a doctor's consulting room, a clinic, a hospital or a health care provider may have visited you at home.

6000. Have you received any health care in the last 12 months?

- Yes..... **1**
- No **5 (Go to 6600)**

6001. In the last 12 months, did you get any health care at an outpatient health facility or did a health care provider visit you at home? An outpatient health facility is a doctor's consulting room, a clinic or a hospital outpatient unit – any place outside your home where you did not stay overnight.

- Yes, at a facility or visited at home..... **1**
- No..... **5 (Go to 6300)**

6002. In the last 12 months, did you get most of your health care at a health facility or most of it from a health provider who visited you in your home?

- Mostly at a health facility..... **1**
- Mostly from a health provider in my home... **2 (Go to 6200)**
- Equally from both**3**

6003. When was your last visit to a health facility or provider? Was it...

- In the last 30 days? **1**
- In the last 3 months? **2**
- In the last 6 months **3**
- Between 6 months and 12 months ago..... **4**
- Don't remember **5**

6004. What was the name of the health care facility?

(Please fill in name of facility, e.g. Oxford Clinic. Only fill in the name of the provider if the facility does not have another name.)

Name: _____

6005. Was [name provided in 6004] your usual place of care?

- Yes..... **1**
- No..... **5**

Go to 6100.

Prompt attention

The next questions are about how promptly you got care.

6100. In the last 12 months, how long did you usually have to wait from the time that you wanted care to the time that you received care?

_____ minutes
 _____ hours
 _____ days
 _____ weeks
 _____ months

6101. In the last 12 months, when you wanted care, how often did you get care as soon as you wanted?

Always..... 1
 Usually 2
 Sometimes..... 3
 Never..... 4

6102. In the last 12 months have you needed any laboratory tests or examinations? Some examples of tests or special examinations are blood tests, scans or X-rays.

Yes..... 1
 No..... 5(Go to 6104)

6103. Generally, how long did you have to wait before you could get the laboratory tests or examinations done?

Got them same day..... 1
 1-2 days..... 2
 3-5 days..... 3
 6-10 days..... 4
 More than 10 days
 (specify)..... 5

6104. Now, overall, how would you rate your experience of getting prompt attention at the health services in the last 12 months? Prompt attention means ... (Read the prompt attention card to the respondent).

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

Dignity

The next questions are about the dignity with which you were treated when you sought health care.

6110. In the last 12 months, when you sought health care, how often did doctors, nurses or other health care providers treat you with respect?

| | |
|----------------|---|
| Always..... | 1 |
| Usually | 2 |
| Sometimes..... | 3 |
| Never | 4 |

6111. In the last 12 months, how often did the office staff, such as receptionists or clerks there, treat you with respect?

| | |
|----------------|---|
| Always..... | 1 |
| Usually | 2 |
| Sometimes..... | 3 |
| Never | 4 |

6112. In the last 12 months, how often were your physical examinations and treatments done in a way that your privacy was respected?

| | |
|----------------|---|
| Always..... | 1 |
| Usually | 2 |
| Sometimes..... | 3 |
| Never | 4 |

6113. Now, overall, how would you rate your experience of getting treated with dignity at the health services in the last 12 months? Dignity means ...
(Read the dignity card to the respondent).

| | |
|----------------|---|
| Very good..... | 1 |
| Good | 2 |
| Moderate | 3 |
| Bad | 4 |
| Very bad..... | 5 |

Communication

The next questions are about how health care providers communicated with you when you sought health care.

6120. In the last 12 months, how often did doctors, nurses or other health care providers listen carefully to you?
- | | |
|-----------------|---|
| Always..... | 1 |
| Usually | 2 |
| Sometimes | 3 |
| Never | 4 |
6121. In the last 12 months, how often did doctors, nurses or other health care providers, explain things in a way you could understand?
- | | |
|-----------------|---|
| Always..... | 1 |
| Usually | 2 |
| Sometimes | 3 |
| Never | 4 |
6122. In the last 12 months, how often did doctors, nurses, or other health care providers give you time to ask questions about your health problem or treatment?
- | | |
|-----------------|---|
| Always..... | 1 |
| Usually | 2 |
| Sometimes | 3 |
| Never | 4 |
6123. Now, overall, how would you rate your experience of how well health care providers communicated with you in the last 12 months? Communication means **(Read the communication card to the respondent)**.
- | | |
|----------------|---|
| Very good..... | 1 |
| Good | 2 |
| Moderate | 3 |
| Bad | 4 |
| Very bad | 5 |

Autonomy

As part of your care, decisions are made about which treatments or tests to give. The next questions are your involvement in decisions about the care and treatment you received in the last 12 months.

6130. In the last 12 months, when you went for health care, were any decisions made about your care, treatment (giving you drugs, for example) or tests?

- Yes..... **1**
- No..... **5(Go to 6132)**

6131. In the last 12 months, how often did doctors, nurses or other health care providers there involve you as much as you wanted be in deciding about the care, treatment or tests?

- Always..... **1**
- Usually **2**
- Sometimes..... **3**
- Never **4**

6132. In the last 12 months, how often did doctors, nurses or other health care providers there ask your permission before starting the treatment or tests?

- Always..... **1**
- Usually **2**
- Sometimes..... **3**
- Never **4**

6133. Now, overall, how would you rate your experience of getting involved in making decisions about your care or treatment as much as you wanted in the last 12 months? Being involved in decision making means ... **(Read the autonomy card to the respondent).**

- Very good..... **1**
- Good..... **2**
- Moderate **3**
- Bad **4**
- Very bad..... **5**

Confidentiality of Information

The next questions are about your experience of confidentiality of information in the health services.

6140. In the last 12 months, how often were talks with your doctor, nurse or other health care provider done privately so other people who you did not want to hear could not overhear what was said?

Always..... 1
Usually 2
Sometimes 3
Never..... 4

6141. In the last 12 months, how often did your doctor, nurse or other health care provider keep your personal information confidential? This means that anyone whom you did not want informed could not find out about your medical conditions.

Always..... 1
Usually 2
Sometimes 3
Never..... 4

6142. Now, overall, how would you rate your experience of the way the health services kept information about you confidential in the last 12 months? Confidentiality of information means ... **(Read the confidentiality card to the respondent).**

Very good..... 1
Good 2
Moderate 3
Bad 4
Very bad 5

Choice

The next questions are about the choice of health care providers you have.

6150. Over the last 12 months, with the doctors, nurses and other health care providers available to you how big a problem, if any, was it to get to a health care provider you were happy with?

- No problem..... 1
- Mild problem..... 2
- Moderate problem 3
- Severe problem..... 4
- Extreme problem..... 5

6151. Over the last 12 months, how big a problem, if any, was it to get to use other health services other than the one you usually went to?

- No problem..... 1
- Mild problem..... 2
- Moderate problem 3
- Severe problem..... 4
- Extreme problem..... 5
- NA – never tried 9

6152. Now, overall, how would you rate your experience of being able to use a health care provider or service of your choice over the last 12 months? Choice means ... **(Read the choice card to the respondent).**

- Very good..... 1
- Good..... 2
- Moderate 3
- Bad 4
- Very bad..... 5

Quality of Surroundings or Environment

The next questions are about the environment or the surroundings at the places you go to for health care.

6160. Thinking about the places you visited for health care in the last 12 months, how would you rate the basic quality of the waiting room, for example, space, seating and fresh air?

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

6161. Thinking about the places you visited for health care over the last 12 months, how would you rate the cleanliness of the place?

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

6162. Now, overall, how would you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanliness of the health services you visited in the last 12 months? Quality of surroundings means ... **(Read the surroundings or environment card to the respondent).**

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

Go to 6300: Inpatient Care

6200. Home Care

Now for all the following questions on health care you receive at home, I would like you to think about all the health care providers who visited you at home over the last 12 months.

The next questions are about how promptly you received care. Sometimes you need care right away for an injury or illness and sometimes you do not need it right away, but can wait for an appointment. The next questions ask about those two different kinds of situations and how promptly you got care.

Prompt Attention

The next questions are about how promptly you got care.

6200. In the last 12 months, how long did you usually have to wait from the time that you wanted care to the time that you received care?

- _____ minutes
- _____ hours
- _____ days
- _____ weeks
- _____ months

6201. In the last 12 months, when you wanted care, how often did you get care as soon as you wanted?

- Always..... 1
- Usually 2
- Sometimes..... 3
- Never 4

6202. In the last 12 months, have you needed any laboratory tests or examinations? Some examples of tests or special examinations are blood tests, scans or X-rays.?

- Yes..... 1
- No..... 5(Go to 6204)

6203. How long did you have to wait before you could get the laboratory tests or examinations done?

- Got them same day..... 1
- 1-2 days..... 2
- 3-5 days..... 3
- 6-10 days..... 4
- More than 10 days
(specify)..... 5

6204. Now, overall, how would you rate your experience of getting prompt attention

at your home in the last 12 months? Prompt attention means ... **(Read the prompt attention card to the respondent).**

- Very good..... 1
- Good 2
- Moderate 3
- Bad 4
- Very bad 5

Dignity

The next questions are about the dignity with which you were treated when you were treated in your home visit.

6210. In the last 12 months, when you were visited at home, how often did doctors, nurses or other health care providers treat you with respect?

- Always..... 1
- Usually 2
- Sometimes 3
- Never 4

6211. In the last 12 months, how often were your physical examinations and treatments conducted during your home visit done in such a way that they ensured that your privacy was respected?

- Always..... 1
- Usually 2
- Sometimes 3
- Never 4

6212. Now, overall, how would you rate your experience of getting treated with dignity by the health services in the last 12 months? Dignity means ... **(Read the dignity card to the respondent).**

- Very good..... 1
- Good 2
- Moderate 3
- Bad 4
- Very bad 5

Communication

The next questions are about how health care providers communicated with you when they visited you at home.

6220. In the last 12 months, how often did doctors, nurses or other health care providers who visited you listen carefully to you?

Always..... 1
 Usually 2
 Sometimes 3
 Never 4

6221. In the last 12 months, how often did doctors, nurses or other health care providers, explain things in a way you could understand?

Always..... 1
 Usually 2
 Sometimes 3
 Never 4

6222. In the last 12 months, how often did doctors, nurses, or other health care providers give you time to ask questions about your health problem or treatment?

Always..... 1
 Usually 2
 Sometimes 3
 Never 4

6223. Now, overall, how would you rate your experience of how well health care providers communicated with you in the last 12 months? Communication means ... **(Read the communication card to the respondent).**

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

Autonomy

As part of your care, decisions are made about which treatments or tests to give. The next questions are your involvement in decisions about the care and treatment you received in the last 12 months.

6230. In the last 12 months, when you were visited at home, were any decisions made about your care, treatment (giving you drugs, for example) or tests?

- Yes..... 1
- No..... 5(Go to 6232)

6231. In the last 12 months, how often did doctors, nurses or other health care providers involve you as much as you wanted be in deciding about the care, treatment or tests?

- Always..... 1
- Usually 2
- Sometimes..... 3
- Never 4

6232. In the last 12 months, how often did doctors, nurses or other health care providers ask your permission before starting the treatment or tests?

- Always..... 1
- Usually 2
- Sometimes..... 3
- Never 4

6233. Now, overall, how would you rate your experience of getting involved in making decisions about your care or treatment as much as you wanted in the last 12 months? Being involved in decision making means ... **(Read the autonomy card to the respondent).**

- Very good..... 1
- Good 2
- Moderate 3
- Bad 4
- Very bad 5

Confidentiality of Information

The next questions are about your experience of confidentiality of information in the health services.

6240. In the last 12 months, how often were talks with your doctor, nurse or other health care provider in your home visits done privately so other people who you did not want to hear could not overhear what was said?

Always..... 1
 Usually 2
 Sometimes 3
 Never..... 4

6241. In the last 12 months, how often did your doctor, nurse or other health care provider keep your personal information confidential? This means that anyone whom you did not want informed could not find out about your medical conditions.

Always..... 1
 Usually 2
 Sometimes 3
 Never..... 4

6242. Now, overall, how would you rate your experience of the way the health services kept information about you confidential in the last 12 months? Confidentiality means ... **(Read the confidentiality of information card to the respondent).**

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

Choice

The next questions are about the choice of health care providers you have.

6250. Over the last 12 months, with the doctors, nurses and other health care providers available to you how big a problem, if any, was it to get to a health care provider you were happy with?

- No problem..... 1
- Mild problem..... 2
- Moderate problem 3
- Severe problem..... 4
- Extreme problem 5

6251. Over the last 12 months, how big a problem, if any, was it to get to use other health services other than the one you usually went to?

- No problem..... 1
- Mild problem..... 2
- Moderate problem 3
- Severe problem..... 4
- Extreme problem 5
- NA 9

6252. Now, overall, how would you rate your experience of being able to use a health care provider or service of your choice over the last 12 months? Choice means ... (Read the choice card to the respondent).

- Very good..... 1
- Good..... 2
- Moderate 3
- Bad 4
- Very bad..... 5

Go to 6300: Inpatient Care

6300. Inpatient Care

Now I would like to ask you some questions about getting health care from a place where you stay over night, which in most cases are hospitals.

6300. Have you stayed overnight in a health care centre or hospital in the last 12 months?

Yes..... 1 No.....5 (Go to 6400)

6301. What was the name of the hospital you stayed in most recently?
(Please fill in name of facility, e.g. Oxford Hospital)

Name: _____

6302. Did you get your hospital care as soon as you wanted?

Yes..... 1
No 5

6303. When you were in the hospital, how often did you get attention from doctors and nurses as quickly as you wanted?

Always..... 1
Usually 2
Sometimes 3
Never 4

6304. Now, overall, how would you rate your experience of getting prompt attention at the hospital in the last 12 months? Prompt attention means ... (Read the **prompt attention card to the respondent**).

Very good..... 1
Good 2
Moderate 3
Bad 4
Very bad 5

6305. Overall, how would you rate your experience of getting treated with dignity at the hospital in the last 12 months? Dignity means ... (Read the **dignity card to the respondent**).

Very good..... 1
Good 2
Moderate 3
Bad 4
Very bad 5

6306. Overall, how would you rate your experience of how well health care providers communicated with you during your stay in the hospital in the last 12 months? Communication means ... **(Read the communication card to the respondent).**

- Very good 1
- Good 2
- Moderate 3
- Bad 4
- Very bad 5

6307. Overall, how would you rate your experience of getting involved in making decisions about your care or treatment as much as you wanted when you were in hospital in the last 12 months? Being involved in decision making means ... **(Read the autonomy card to the respondent).**

- Very good 1
- Good 2
- Moderate 3
- Bad 4
- Very bad 5

6308. Overall, how would you rate your experience of the way the hospital kept personal information about you confidential in the last 12 months. Confidentiality means ... **(Read the confidentiality of information card to the respondent).**

- Very good 1
- Good 2
- Moderate 3
- Bad 4
- Very bad 5
- DK 8

6309. Overall, how would you rate your experience of being able to use a hospital of your choice over the last 12 months? Choice means ... **(Read the choice card to the respondent).**

- Very good 1
- Good 2
- Moderate 3
- Bad 4
- Very bad 5

6310. Overall, how would you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanliness of the health services you visited in the last 12 months? Quality of surroundings means ... **(Read the surroundings or environment card to the respondent).**

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

6311. In the last 12 months, when you stayed in a hospital, how big a problem, if any, was it to get the hospital to allow your family and friends to take care of your personal needs, such as bringing you your favourite food, soap etc..?

No problem..... 1
 Mild problem..... 2
 Moderate problem 3
 Severe problem..... 4
 Extreme problem..... 5

6312. During your stay in the hospital, how big a problem, if any, was it to have the hospital allow you to practice religious or traditional observances if you wanted to? Would you say it was:

No problem..... 1
 Mild problem..... 2
 Moderate problem..... 3
 Severe problem..... 4
 Extreme problem..... 5

6313. Now, overall, how would you rate your experience of how the hospital allowed you to interact with family, friends and to continue your social and/ or religious customs during your stay over the last 12 months? Social support means ... **(Read the social support card to the respondent).**

Very good..... 1
 Good 2
 Moderate 3
 Bad 4
 Very bad 5

Go to 6400: Other Aspects of the Health System

6400. Other Aspects of the Health System

6400. In the last 12 months were you treated badly by the health system or services in your country because of your: **(Check all that apply)**

| | Yes | No | Refuse |
|------------------------------|-----|----|--------|
| 1. Nationality | 1 | 5 | 7 |
| 2. Social class | 1 | 5 | 7 |
| 3. Lack of private insurance | 1 | 5 | 7 |
| 4. Ethnicity | 1 | 5 | 7 |
| 5. Colour | 1 | 5 | 7 |
| 6. Sex | 1 | 5 | 7 |
| 7. Language | 1 | 5 | 7 |
| 8. Religion | 1 | 5 | 7 |
| 9. Political/other beliefs | 1 | 5 | 7 |
| 10. Health status | 1 | 5 | 7 |
| 11. Lack of wealth | 1 | 5 | 7 |
| 12. Other (specify) _____ | 1 | 5 | 7 |

This question is only to be asked to women.

6401. In the last 12 months when you used health services in this country, did you feel that you were treated worse because you were a woman?

Yes..... 1
 No..... 5
 Refuse..... 7

6500. Utilization

I will read you a list of different types of places you can get health services. Please can you indicate the number of times you went to each of them in the last 30 days.

Times

- 6500 _____ General Practitioners (doctors)
- 6501 _____ Dentists
- 6502 _____ Specialists
- 6503 _____ Chiropractors
- 6504 _____ Traditional Healers
- 6505 _____ Clinics (staffed mainly by nurses, run separately from hospital)
- 6506 _____ Hospital outpatient unit
- 6507 _____ Hospital inpatient services
- 6508 _____ Pharmacy (where you talked to someone about your care and did not only
purchase medicine)
- 6509 _____ Home health care services
- 6510 _____ Other (specify) _____

6511. What was the main reason that you went to the health care provider for your most recent visit? I will read through a list. Please indicate all that apply. **(Check all that apply).**

| | Yes | No | DK | NA |
|---|-----|----|----|----|
| 1. You needed a check up for a chronic, ongoing problem | 1 | 5 | 8 | 9 |
| 2. You needed care because my chronic, ongoing problem flared up | 1 | 5 | 8 | 9 |
| 3. You needed care because of an injury or illness that had just happened | 1 | 5 | 8 | 9 |
| 4. You needed to follow up with the provider after having an operation or treatment for an injury | 1 | 5 | 8 | 9 |
| 5. You were not sick, you went for a general exam or preventive care | 1 | 5 | 8 | 9 |
| 6. Other (specify) _____ | 1 | 5 | 8 | 9 |

6512. What services were provided at your most recent visit? Again, I will read through a list. Please indicate all that apply **(Check all that apply).**

| | Yes | No | DK | NA |
|--|-----|----|----|----|
| 1. You were examined | 1 | 5 | 8 | 9 |
| 2. You received tests | 1 | 5 | 8 | 9 |
| 3. The health care provider gave you treatment | 1 | 5 | 8 | 9 |
| 4. The health care provider talked with you about your health problem | 1 | 5 | 8 | 9 |
| 5. The health care provider talked to you about your health in general | 1 | 5 | 8 | 9 |
| 6. You picked up medicine or a prescription | 1 | 5 | 8 | 9 |
| 7. Other (specify) _____ | 1 | 5 | 8 | 9 |

Go to 6600: Review of Health System

6600. Review of Health System

6600. In the last 12 months, were you ever refused health care because you could not afford it?

Yes..... 1
No..... 5

6601. In the last 12 months, did you not seek health care because you could not afford it?

Yes..... 1
No..... 5

Ask the respondent to read the cards below or read the cards to the respondent if he/she would prefer. These are descriptions of some different ways the health care services in your country show respect for people and make them the centre of care. Please write the code in the space provided.

Thinking about what is on these cards and about the whole health system, which is the most important and the least important to you?

6602. **MOST IMPORTANT** _____ ¹Most important

6603. **LEAST IMPORTANT** _____ ⁸Least important

| | |
|---|-------------------|
| DIGNITY | CODE = DIG |
| <ul style="list-style-type: none">◆ being shown respect◆ having physical examinations conducted in privacy | |

| | |
|---|------------------|
| CONFIDENTIALITY OF INFORMATION | CODE = CI |
| <ul style="list-style-type: none">◆ having your medical history kept confidential◆ having talks with health providers done so that other people who you don't want to have hear you can't overhear you | |

| | |
|---|------------------|
| CHOICE | CODE = CH |
| <ul style="list-style-type: none">◆ being able to choose your doctor or nurse or other person usually providing your health care◆ being able to go to another place for health care if you want to | |

| | |
|--|------------------|
| PROMPT ATTENTION | CODE = PA |
| <ul style="list-style-type: none">◆ there is a reasonable distance and travel time from your home to the health care provider◆ you get fast care in emergencies◆ you have short waiting times for appointments and consultations, and get tests done quickly◆ short waiting lists for non-emergency surgery | |

| | |
|--|-------------------|
| AUTONOMY | CODE = AUT |
| <ul style="list-style-type: none"> ◆ being involved in deciding on your care or treatment if you want to ◆ having the provider ask your permission before starting treatments or tests | |

| | |
|---|-------------------|
| SURROUNDINGS OR ENVIRONMENT | CODE = ENV |
| <ul style="list-style-type: none"> ◆ having enough space, seating and fresh air in the waiting room ◆ having a clean facility (including clean toilets) ◆ having healthy and edible food | |

| | |
|--|------------------|
| SOCIAL SUPPORT | CODE = SS |
| <ul style="list-style-type: none"> ◆ the provision of food and other gifts by relatives ◆ freedom of religious practices | |

| | |
|---|-------------------|
| COMMUNICATION | CODE = COM |
| <ul style="list-style-type: none"> ◆ the provider listens to you carefully ◆ the provider explains things so you can understand ◆ you have time to ask questions | |

6604. Did the respondent read the cards him/herself? Yes....1 / No.....5

Annex B2. World Health Survey Questionnaire on Responsiveness

7000. Health System Responsiveness

Time Begin: ____ : ____

Needing Health Care And General Evaluation Of Health Systems

| | | | |
|-------|---|---|---------------------|
| Q7000 | When was the last time that either you as an adult, or a child of yours aged 12 years or less, needed health care? [Interviewer: stop reading further as soon as the respondent has selected one.] | 1. In the last 30 days 2. Between 1 month and less than 1 year ago 3. Between 1 year and less than 2 years ago 4. Between 2 years and less than 3 years ago 5. Between 3 years and less than 5 years ago 6. More than 5 years ago 7. Never needed | If 7, Go to Q7020 |
| Q7001 | Was the <u>last</u> need for health care for yourself or for your child? [Interviewer: Use "you" or "your child" according to the response] | 1. Yourself 2. Your child | |
| Q7002 | Thinking of the last time you [your child] needed to see a health care provider who could treat your condition, how many were there around who you could chose from? Interviewer: RECORD NUMBER | | |
| Q7003 | Which reason <u>best</u> describes why you [your child] last needed health care? [Interviewer - the respondent may select ONLY one] | 1. High fever, severe diarrhea, or cough 2. Immunization 3. Antenatal consultation 4. Family planning 5. Childbirth 6. Dental care 7. Arthritis 8. Asthma 9. Heart disease 10. Bodily injury 11. Minor surgery 12. Other | |
| Q7004 | The last time you [your child] needed health care, did you get health care? | 1. Yes 5. No | If Yes, Go to Q7016 |

Which reasons best explain why you [your child] did not get health care?

| | | | |
|-------|---|---|---|
| Q7005 | Could not afford the cost of the visit | 1. Yes | 5. No |
| Q7006 | No transport | 1. Yes | 5. No |
| Q7007 | Could not afford the cost of transport | 1. Yes | 5. No |
| Q7008 | The health care provider's drugs or equipment are inadequate | 1. Yes | 5. No |
| Q7009 | The health care provider's skills are inadequate | 1. Yes | 5. No |
| Q7010 | You were previously badly treated | 1. Yes | 5. No |
| Q7011 | Could not take time off work or had other commitments | 1. Yes | 5. No |
| Q7012 | You did not know where to go | 1. Yes | 5. No |
| Q7013 | You thought you were not sick enough | 1. Yes | 5. No |
| Q7014 | You tried but were denied health care | 1. Yes | 5. No |
| Q7015 | Other | 1. Yes | 5. No |
| Q7016 | When you last needed health care, where did you get care? | 1. At a health care provider, excluding an overnight stay in hospital 2. At a hospital where you stayed overnight 3. At home | |
| Q7017 | The last time you [your child] sought care for [refer to the CONDITIONS listed in Q 7002] did the health care provider prescribe any medicine for you [your child]? | 1. Yes | 5. No |
| Q7018 | Of the medicines that were prescribed for you [your child], how many of them were you able to get? | 1. All of them | 8. DK |
| Q7019 | Which reason best explains why you [your child] did not get all the medicines you were prescribed? | 1. Could not afford 2. Could not find all medicines 3. Did not believe all the medications were needed 4. Started to feel better 5. Already had some of the medicines at home 6. Other | |
| | | | If No or DK, Go to Q7020 If All of them, Go to Q7020 |

| | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
|---|---|---|--------------------------------------|------------------------|----------------------|
| Q7020 How would you rate the way health care in your country involves you in deciding what services it provides and where it provides them? | 1. Very satisfied | 2. Fairly satisfied | 3. Neither satisfied or dissatisfied | 4. Fairly dissatisfied | 5. Very dissatisfied |
| Q7021 In general would you say you are very satisfied, fairly satisfied, neither satisfied nor dissatisfied, fairly dissatisfied or very dissatisfied with the way health care runs in your country? | 1. Yes, for a person living in the same household | 2. Yes, for a person living in a separate household | 3. No | | |
| Q7022 During the past year, did you provide help to a relative or friend (adult or child), because this person has a long-term physical or mental illness or disability, or is getting old and weak? | Please tell me the kind of care you provided : | | | | |
| Q7023 You helped with personal care, such as going to the toilet, washing, getting dressed, or eating | 1. Yes | | 5. No | | |
| Q7024 You helped with medical care, like changing bandages and giving medicines | 1. Yes | | 5. No | | |
| Q7025 You helped with household activities, such as meal preparation, shopping, cleaning, laundry | 1. Yes | | 5. No | | |
| Q7026 You watched over them since their behaviour can be upsetting or dangerous to themselves or others | 1. Yes | | 5. No | | |
| Q7027 You helped them to get around outside the home | 1. Yes | | 5. No | | |
| In your dealings with private health care organizations or the government, have you ever had any difficulties: | | | | | |
| Q7028 Obtaining payment exemptions or the right to special rates for health care | 1. Yes | 5. No | | 9. Not applicable | |
| Q7029 Completing or filling out applications for health insurance | 1. Yes | 5. No | | 9. Not applicable | |
| Q7030 Finding out what benefits you are entitled to from your health insurance | 1. Yes | 5. No | | 9. Not applicable | |
| Q7031 Getting reimbursements from health insurance organizations | 1. Yes | 5. No | | 9. Not applicable | |

If No, Go to Q7028

IMPORTANCE

| | | | | | |
|---|--------------------------------------|---------------------------------|---------------------------------------|-------------------------------------|--------------------------------|
| <p>Q7100 How important is "respectful treatment" to you. This means</p> <ul style="list-style-type: none"> • being shown respect when greeted by and when talking to health care providers •having physical examinations conducted in a way that respects your cultural norms <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |
| <p>Q7101 How important is "confidentiality of personal information" to you. This means</p> <ul style="list-style-type: none"> •having information about your health and other personal information kept confidential •having conversations with health care providers without other people overhearing <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |
| <p>Q7102 How important is "convenient travel and short waiting times" to you. This means</p> <ul style="list-style-type: none"> •having short travel times and convenient access to health care facilities •having short waiting times for consultations and hospital admissions <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |

| | | | | | |
|--|--------------------------------------|---------------------------------|---------------------------------------|-------------------------------------|--------------------------------|
| <p>Q7103</p> <p>How important is "choice of health care providers" to you. This means,</p> <ul style="list-style-type: none"> •being able to choose your health care provider (place or person) •being able to consult for a second opinion or with a specialist if so desired <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |
| <p>Q7104</p> <p>How important is "involvement in decision making" to you. This means</p> <ul style="list-style-type: none"> •being involved as much as you want in deciding about your health care •freedom to discuss other treatment options or care regimes if you want <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |
| <p>Q7105</p> <p>How important are "good quality surroundings" to you? This means</p> <ul style="list-style-type: none"> •having enough space, seating and fresh air in the waiting rooms, examination rooms and hospital wards •having a clean facility (including clean toilets) <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |

| | | | | | |
|--|-------------------------------|--------------------------|--------------------------------|------------------------------|-------------------------|
| <p>Q7106 How important is "contact with the outside world " to you? This means</p> <ul style="list-style-type: none"> •having family and friends visit you as much as you want when you are a patient in hospital •being able to keep in contact with family and friends and to have information about what is happening outside the hospital <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |
| <p>Q7107 How important is "clarity of communication" to you. This means</p> <ul style="list-style-type: none"> • having the health care providers explain things in a way you can understand • having enough time to ask questions if you don't understand something <p>Would you say it is:</p> | <p>1. Extremely Important</p> | <p>2. Very Important</p> | <p>3. Moderately Important</p> | <p>4. Slightly Important</p> | <p>5. Not Important</p> |

Seeing Health Care Providers

| | | | |
|--|---|---|---------------------|
| [Interviewer: If an adult went for health care at the same time as for his/her children, focus on the adult's experience] | | | |
| Q7200 | Over the last <u>5</u> years, was there ever a time you stayed overnight in a hospital or other type of long term care facility for your own health care? | 1. Yes - a hospital facility 2. Yes - long term care facility 5. No | IF Yes, Go to Q7400 |
| Q7201 | Over the last <u>5</u> years, was there ever a time that one of your children aged <u>12</u> years or less stayed overnight in a hospital? [Interviewer: if the person has more than one child, ask for sex and age of the child that had the last visit; If several children were seen at the same time, focus on the YOUNGEST child.] | 1. Yes 5. No | IF No, Go to Q7204 |
| Q7202 | What is the <u>sex</u> of the child? | 1. Female 2. Male | |
| Q7203 | What is the date of <u>birth</u> of the child? | MM YY | |
| Q7204 | Over the last <u>12</u> months, did you receive any health care excluding any overnight stay in hospital? | 1. Yes 5. No | IF Yes, Go To Q7300 |
| Q7205 | Over the last <u>12</u> months, was there ever a time you accompanied one of your children aged <u>12</u> years or less for health care excluding any overnight stay in hospital? [Interviewer: if the person has more than one child ask for sex and age of the child that had the last visit; If several children were seen at the same time, focus on the YOUNGEST child.] | 1. Yes 5. No | IF No, Go To Q8000 |
| Q7206 | What is the <u>sex</u> of the child? | 1. Female 2. Male | |
| Q7207 | What is the date of <u>birth</u> of the child? | MM YY | |

Outpatient and Care at Home

[Interviewer: use "your" or "your child" according to responses in previous section "Seeing Health Care Providers".]

| | | |
|--------------|--|--|
| Q7300 | What was the name of the last health care provider you [your child] used in the last 12 months? [Interviewer: try get the name of the clinic or health centre, rather than the doctor. If the respondent used a clinic or health centre. If the respondent was visited at home, write "home visit".] | |
| Q7301 | Was the last place you [your child] visited in the last 12 months: | 1. operated by the government 2. privately operated 3. other |
| Q7302 | Which was the last health care provider you visited? [Interviewer: After q7302 substitute the type of health care provider selected by the patient when you see [health care provider] in parentheses.] | 1. Medical doctor (including gynaecologist, psychiatrist, ophthalmologist, etc.) 2. Nurse 3. Midwife 4. Dentist 5. Physiotherapist or chiropractor 6. Traditional medicine practitioner 7. Other |
| Q7303 | What was the sex of [the health care provider]? | 1. Female 2. Male |
| Q7304 | In your opinion, was the [health care providers] skill adequate for your [child's] treatment? | 1. Yes 5. No |
| Q7305 | In your opinion, was [the health care providers] equipment adequate for your [child's] treatment? | 1. Yes 5. No |
| Q7306 | In your opinion, were [the health care providers] drug supplies adequate for your [child's] treatment? | 1. Yes 5. No |
| Q7307 | Thinking about your [child's] last visit, how long did it take you to get there? (minutes) | |
| Q7308 | Thinking about your [child's] last visit, how did you get there? [Interviewer: mark the one used for most of the travel distance.] | 1. Private car 2. Public or motorcycle / transport 3. Ambulance 4. Bicycle 5. Walked 6. Other |

Thinking about your [child's] last visit, how much did you or your household pay for (local currency): [Interviewer: only write 0 if the service was free. If a person did not have tests or drugs, circle "Not applicable, not have"]

| | | |
|--------------|------------------------------|--|
| Q7309 | [Health care providers] fees | 8. DK |
| Q7310 | Medicines | 8. DK 9. Not Applicable, Not have |

| | | | | | |
|-------|---|--------------|---------|-------------|--------------------------------------|
| Q7311 | Tests | | 8. DK | | 9. Not applicable, not have |
| Q7312 | Transport | | 8. DK | | 9. Not applicable, not have |
| Q7313 | Other | | 8. DK | | 9. Not applicable, not have |
| Q7314 | Did you or your household pay less than the normal health care fees because of a government discount or exemption? | 1. Yes | 8. DK | | 9. Not applicable, not have |
| Q7315 | For your [child's] last visit, how would you rate the travelling time to [the health care provider]? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7316 | For your [child's] last visit, how would you rate the amount of time you waited before being attended to? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7317 | For your [child's] last visit, how would you rate your experience of being greeted and talked to respectfully? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7318 | For your [child's] last visit, how would you rate the way your privacy was respected during physical examinations and treatments? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7319 | For your [child's] last visit, how would you rate the experience of how clearly health care providers explained things to you? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7320 | For your [child's] last visit, how would you rate your experience of getting enough time to ask questions about your health problem or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7321 | For your [child's] last visit, how would you rate your experience of getting information about other types of treatments or tests? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7322 | For your [child's] last visit, how would you rate your experience of being involved in making decisions about your health care or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| Q7323 | For your [child's] last visit, how would you rate the way the health services ensured you could talk privately to health care providers? | 1. Very good | 2. Good | 3. Moderate | 4. Bad |
| | | | | | 9. Not applicable, no exam/treatment |

| | | | | | | | |
|--------------|--|--------------|---------|-------------|--------|-------------|------------------------------|
| Q7324 | For your [child's] last visit, how would you rate the way your personal information was kept confidential? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad | 8. DK |
| Q7325 | For your [child's] last visit, how would you rate the freedom you had to choose your [health care provider]? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad | |
| Q7326 | For your [child's] last visit, how would you rate the cleanliness of the rooms inside the facility, including toilets? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad | 9. Not applicable, home care |
| Q7327 | For your [child's] last visit, how would you rate the amount of space in the waiting and examination rooms? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad | 9. Not available, home care |

In the last 12 months did you feel that you were treated worse by health care providers for any of the following reasons. Because of your:

| | | | |
|--------------|------------------------|--------|-------|
| Q7328 | Sex | 1. Yes | 5. No |
| Q7329 | Age | 1. Yes | 5. No |
| Q7330 | Lack of money | 1. Yes | 5. No |
| Q7331 | Social class | 1. Yes | 5. No |
| Q7332 | Ethnic group or colour | 1. Yes | 5. No |
| Q7333 | Type of illness | 1. Yes | 5. No |
| Q7334 | Nationality | 1. Yes | 5. No |

Inpatient Hospital

| | | | | | |
|--------------|---|--|-----------------------|----------------------|----------------------|
| | [Interviewer: use "you" or "your child" according to responses in previous section "Seeing Health Care Providers".] | | | | |
| Q7400 | What was the name of the last hospital or long term care facility you [your child] stayed in, in the last 5 years? | | | | |
| Q7401 | Was the hospital (or long term care facility): | 1. operated by the government | 2. privately operated | 3. other | |
| Q7402 | When was your [child's] last overnight stay? [Interviewer: stop reading further as soon as the respondent has selected one] | 1. In the last weeks | 2. In the last year | 3. In the last years | 4. In the last years |
| Q7403 | Which of the following best describes the reason for your [child's] last overnight stay? | 1. High fever, severe diarrhea, or cough 2. Childbirth 3. Arthritis 4. Asthma 5. Heart disease 6. Bodily injury 7. Minor surgery 8. Other | | | |
| Q7404 | How long was your [child's] stay on this occasion? [Interviewer: stop reading further as soon as the respondent has selected one] | 1. 1-2 days | 2. 3-5 days | 3. 6-14 days | 4. 15 days and more |
| Q7405 | In your opinion, was the skill of the health care providers adequate for your [child's] treatment? | 1. Yes 5. No | | | |
| Q7406 | In your opinion, was the hospital's equipment adequate for your [child's] treatment? | 1. Yes 5. No | | | |
| Q7407 | In your opinion, were the hospital's drug supplies adequate for your [child's] treatment? | 1. Yes 5. No 8. NA | | | |
| Q7408 | Thinking about your [child's] last hospital stay, how long did it take you to get there (in minutes)? | | | | |
| Q7409 | Thinking about your [child's] last hospital stay, how did you get there? [Interviewer, mark the one used for most of the travel distance.] | 1. Private car or motorcycle | 2. Public transport | 3. Ambulance | 4. Bicycle |
| | | | | 5. Walked | 6. Other |

| | | | | | | |
|--|--|--------------|---------------------|----------------------|-----------------------|----------------------|
| Q7410 | For your [child's] last hospital stay, how long from the time you needed hospital care did you wait to be admitted to hospital? [Interviewer: stop reading further as soon as the respondent has selected one] | 1. Same day | 2. Less than a week | 3. Less than 1 month | 4. Less than 3 months | 5. 3 months and more |
| Thinking about your last hospital stay, how much did you or your household pay for (local currency): | | | | | | |
| Q7411 | doctor's fees | | | | | -8, DK |
| Q7412 | Drugs | | | | | -8, DK |
| Q7413 | Tests | | | | | -8, DK |
| Q7414 | Transport | | | | | -8, DK |
| Q7415 | Other | | | | | -8, DK |
| Q7416 | Did you or your household pay less than the normal health care fees because of a government discount or exemption? | 1. Yes | 5. No | | | 9. NA |
| Q7417 | Thinking about your [child's] last hospital stay, how many people slept in the same room as you [your child]? | | | | | |
| Q7418 | For your [child's] last hospital stay, how would you rate the travelling time to the hospital? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7419 | For your [child's] last hospital stay, how would you rate the amount of time you waited before being attended to? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7420 | For your [child's] last hospital stay, how would you rate your experience of being greeted and talked to respectfully? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7421 | For your [child's] last hospital stay, how would you rate the way your [child's] privacy was respected during physical examinations and treatments? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7422 | For your [child's] last hospital stay, how would you rate the experience of how clearly health care providers explained things to you? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7423 | For your [child's] last hospital stay, how would you rate your experience of getting enough time to ask questions about your [child's] health problem or treatment? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7424 | For your [child's] last hospital stay, how would you rate your experience of getting information about other types of treatments or tests? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7425 | For your [child's] last hospital stay, how would you rate your experience of being involved in making decisions about your [child's] health care or treatment? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |

| | | | | | | |
|-------|--|--------------|---------|-------------|--------|-------------|
| Q7426 | For your [child's] last hospital stay, how would you rate the way the health services ensured you could talk privately to health care providers? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7427 | For your [child's] last hospital stay, how would you rate the way your [child's] personal information was kept confidential? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7428 | For your [child's] last hospital stay, how would you rate the freedom you had to choose the health care providers that attended to you [your child]? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7429 | For your [child's] last hospital stay, how would you rate the cleanliness of the rooms inside the facility, including toilets? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7430 | For your [child's] last hospital stay, how would you rate the amount of space you [your child] had? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7431 | For your [child's] last hospital stay, how would you rate the ease of having family and friends visit you [your child]? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7432 | For your [child's] last hospital stay, how would you rate your [child's] experience of staying in contact with the outside world when you [your child] were in hospital? | 1. Very Good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |

In the last 5 years did you feel that you were treated worse by the health care providers at the hospital for any of the following reasons. Because of your:

| | | | |
|-------|------------------------|--------|-------|
| Q7433 | Sex | 1. Yes | 5. No |
| Q7434 | Age | 1. Yes | 5. No |
| Q7435 | Lack of money | 1. Yes | 5. No |
| Q7436 | Social class | 1. Yes | 5. No |
| Q7437 | Ethnic group or colour | 1. Yes | 5. No |
| Q7438 | Type of illness | 1. Yes | 5. No |
| Q7439 | Nationality | 1. Yes | 5. No |

VIGNETTES FOR HEALTH SYSTEM RESPONSIVENESS

| | |
|-------|--------------------------|
| Q7500 | RECORD SET (A, B, C, D): |
|-------|--------------------------|

I am now going to read you stories about people's experiences with health care services. I want you to think about these people's experiences as if they were your own. Once I have finished reading each story, I will ask you to rate what happened in the story as very good, good, moderate, bad or very bad.

| | | | | | | |
|-------------------|------------|--------------|---------|-------------|--------|-------------|
| Vignette 1 | | | | | | |
| Q7501 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7502 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 2 | | | | | | |
| Q7503 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7504 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 3 | | | | | | |
| Q7505 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7506 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 4 | | | | | | |
| Q7507 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7508 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 5 | | | | | | |
| Q7509 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7510 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 6 | | | | | | |
| Q7511 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7512 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |

| | | | | | | |
|--------------------|------------|--------------|---------|-------------|--------|-------------|
| Vignette 7 | | | | | | |
| Q7513 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7514 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 8 | | | | | | |
| Q7515 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7516 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 9 | | | | | | |
| Q7517 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7518 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Vignette 10 | | | | | | |
| Q7519 | Question 1 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |
| Q7520 | Question 2 | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very Bad |

Time End: ____ : ____ : ____

Responsiveness vignettes. Set A

I am now going to read you stories about people's experiences with health care services. I want you to think about these people's experiences as if they were your own. Once I have finished reading each story, I will ask you to rate what happened in the story as very good, good, moderate, bad or very bad.

Respectful Treatment and Prompt Attention

| | | | | | | |
|---|---|--------------|---------|-------------|--------|-------------|
| [Niels] woke up with a sore back so he decided to go to the clinic. It took 30 minutes to travel to the clinic and he was seen within 5 minutes. | | | | | | |
| Q7501 | How would you rate his travelling time? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7502 | How would you rate the amount of time he waited before being attended to? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Anya] took her baby for a vaccination. The nurse said hello and but did not ask for [Anya's] or the baby's name. The nurse also examined [Anya] and made her remove her shirt in the waiting room. | | | | | | |
| Q7503 | How would you rate her experience of being greeted and talked to respectfully? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7504 | How would you rate the way her privacy was respected during physical examinations and treatments? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Stan] broke his leg. It took an hour to be driven to the nearest hospital. He was in pain but had to wait an hour for the surgeon and was only operated on the next day. | | | | | | |
| Q7505 | How would you rate his travelling time? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7506 | How would you rate the amount of time he waited before being attended to? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Conrad] had bad flu. He went to the clinic. The nurse expressed concern about [Conrad's] cough and called the doctor, who gave [Conrad] a full chest examination behind a large screen that hid him from the view of other patients. | | | | | | |
| Q7507 | How would you rate his experience of being greeted and talked to respectfully? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7508 | How would you rate the way his privacy was respected during physical examinations and treatments? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Aitor] broke his hand. He walked to the clinic because there was no transport and the journey took him several hours. He had to sleep near the clinic for two days before seeing the doctor. | | | | | | |
| Q7509 | How would you rate his travelling time? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

| | | | | | | |
|--|---|--------------|---------|-------------|--------|-------------|
| Q7510 | How would you rate the amount of time he waited before being attended to? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Julia] was pregnant and went to the hospital coughing blood. A nurse welcomed her gently and helped her to a private room. A female doctor came to examine her and gave her a clean gown to replace her blood-stained clothes. | | | | | | |
| Q7511 | How would you rate her experience of being greeted and talked to respectfully? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7512 | How would you rate the way her privacy was respected during physical examinations and treatments? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Carla]'s child became seriously sick. [Carla] called an ambulance that arrived after 10 minutes and within 5 minutes they were at the hospital and the doctors were treating the child. | | | | | | |
| Q7513 | How would you rate her travelling time? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7514 | How would you rate the amount of time she waited before being attended to? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Patricia] went to a crowded clinic. At first, no-one greeted her but after waiting for 5 minutes a nurse called her to the examination area where she was examined behind a small screen that mostly hid her from the other patients. | | | | | | |
| Q7515 | How would you rate her experience of being greeted and talked to respectfully? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7516 | How would you rate the way her privacy was respected during physical examinations and treatments? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Karen] needed a blood test. It took her 45 minutes by bus to get to the clinic and she waited for 30 minutes to see the nurse. | | | | | | |
| Q7517 | How would you rate her travelling time? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7518 | How would you rate the amount of time she waited before being attended to? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

[Said] has AIDS. When he goes to his health centre the nurses do not talk to him and deliberately ignore him. During examinations, his clothes are removed and he is made to wait, half-naked in the waiting room.

| | | | | | | |
|-------|---|--------------|---------|-------------|--------|-------------|
| Q7519 | How would you rate his experience of being greeted and talked to respectfully? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7520 | How would you rate the way his privacy was respected during physical examinations and treatments? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

Responsiveness Vignettes. Set B

I am now going to read you stories about people's experiences with health care services. I want you to think about these people's experiences as if they were your own. Once I have finished reading each story, I will ask you to rate what happened in the story as very good, good, moderate, bad or very bad.

Communication and Quality of Basic Amenities

[Thomas] couldn't see well so he went to the doctor and explained the problem. [Thomas] had time to ask the doctor some questions, which the doctor answered until [Thomas] understood almost everything.

| | | | | | | |
|-------|--|--------------|---------|-------------|--------|-------------|
| Q7501 | How would you rate his experience of how clearly health care providers explained things to him? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7502 | How would you rate his experience of getting enough time to ask questions about his health problem or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

[Wing] had his own room in the hospital and shared a bathroom with two others. The room and bathroom were cleaned frequently and had fresh air.

| | | | | | | |
|-------|---|--------------|---------|-------------|--------|-------------|
| Q7503 | How would you rate the cleanliness of the rooms inside the facility, including toilets? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7504 | How would you rate the amount of space [Wing] had? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

[Pia] went to the emergency clinic with a stomach pain. The doctor explained to [Pia] her condition and the treatment. [Pia] asked him some questions and the doctor explained things using examples that were familiar to her until she understood everything.

| | | | | | | |
|-------|--|--------------|---------|-------------|--------|-------------|
| Q7505 | How would you rate her experience of how clearly health care providers explained things to her? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7506 | How would you rate her experience of getting enough time to ask questions about her health problem or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

| | | | | | | |
|-------|--|--------------|---------|-------------|--------|-------------|
| | [Shedra] was hospitalized last year for a hip operation. Her private room had its own bathroom and was comfortable and spacious. It was cleaned by the hospital staff daily. The bed was comfortable and the sheets were changed daily. | | | | | |
| Q7507 | How would you rate the cleanliness of the rooms inside the facility, including toilets? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7508 | How would you rate the amount of space [Shedra] had? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | [Akiko] is in hospital after a car accident. She has lots of scratches, bruises and some broken bones. When the doctor visited her he asked to see her medical records. He asked the nurse some questions and then he said that [Akiko] was making good progress. [Akiko] supposes that she will still stay there for another week but is unsure. | | | | | |
| Q7509 | How would you rate her experience of how clearly health care providers explained things to her? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7510 | How would you rate her experience of getting enough time to ask questions about her health problem or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | [Kamal] has a nervous breakdown and had to spend 3 months in the past year in the local hospital. He had to sleep on an uncomfortable mattress with no sheets. There were 30 other patients in the same dormitory style ward and the toilets would smell, because they were not cleaned. He came back with a skin infection, because he couldn't wash regularly and there were insects in the bed. | | | | | |
| Q7511 | How would you rate the cleanliness of the rooms inside the facility, including toilets? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7512 | How would you rate the amount of space [Kamal] had? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | [Rose] cannot write or read. She went to the doctor because she was feeling dizzy. The doctor didn't have time to answer her questions or to explain anything. He sent her away with a piece of paper without telling her what it said. | | | | | |
| Q7513 | How would you rate her experience of how clearly health care providers explained things to her? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7514 | How would you rate her experience of getting enough time to ask questions about her health problem or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | [Hans] shared his small hospital room with five other patients with no partitions between beds. The five patients shared a toilet, which was cleaned every second day and sometimes smelt. | | | | | |
| Q7515 | How would you rate the cleanliness of the rooms inside the facility, including toilets? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7516 | How would you rate the amount of space [Hans] had? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

[Mario] has been told that he has epilepsy and that he needs to take medication. The doctor has very briefly explained what the condition is. He is very busy and there is a queue of patients waiting to see him. Mario would like to know more about what he has, but feels that there is no time to ask questions. The doctor says goodbye to Mario, and Mario leaves the office.

| | | | | | | |
|--------------|--|--------------|---------|-------------|--------|-------------|
| Q7517 | How would you rate his experience of how clearly health care providers explained things to him? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7518 | How would you rate his experience of getting enough time to ask questions about his health problem or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

[José] shared a hospital room with four other persons. There was a toilet for his ward located along the outside corridor. The room was only occasionally dusty and hot, but without any fan.

| | | | | | | |
|--------------|---|--------------|---------|-------------|--------|-------------|
| Q7519 | How would you rate the cleanliness of the rooms inside the facility, including toilets? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7520 | How would you rate the amount of space [José] had? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

Responsiveness Vignettes. Set C

I am now going to read you stories about people's experiences with health care services. I want you to think about these people's experiences as if they were your own. Once I have finished reading each story, I will ask you to rate what happened in the story as very good, good, moderate, bad or very bad.

Confidentiality, Choice and Involvement

[Simon] was speaking to his doctor about an embarrassing problem. There was a friend and a neighbour of his in the crowded waiting room and because of the noise the doctor had to shout when telling [Simon] the treatment he needed.

| | | | | | | |
|--------------|---|--------------|---------|-------------|--------|-------------|
| Q7501 | How would you rate the way the health services ensured [Simon] could talk privately to health care providers? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7502 | How would you rate the way [Simon's] personal information was kept confidential? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

When the clinic is not busy, [Mamadou] can choose which doctor he sees. But most often it is busy and then he gets sent to whoever is free.

| | | | | | | |
|--------------|--|--------------|---------|-------------|--------|-------------|
| Q7503 | How would you rate [Mamadou's] freedom to choose her health care provider? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
|--------------|--|--------------|---------|-------------|--------|-------------|

The government wanted to expand the town's clinic service. They held one meeting to discuss the expansion plans. [Chad] attended the meeting and freely expressed his views. [Chad] then heard nothing more about what decision the government had made.

| | | | | | | |
|--------------|--|--------------|---------|-------------|--------|-------------|
| Q7504 | How would you rate the way health care involves [Chad] in deciding what services it provides and where it provides them? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
|--------------|--|--------------|---------|-------------|--------|-------------|

| | | | | | | |
|--------------|---|--------------|---------|-------------|--------|-------------|
| | [Rebecca] usually spoke to her doctor about her illness in complete privacy. Once [Rebecca] heard that the doctor spoke to [Rebecca's] friend about her illness. She asked her doctor not to do it again. He did not do it again. | | | | | |
| Q7505 | How would you rate the way the health services ensured [Rebecca] could talk privately to health care providers? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7506 | How would you rate the way [Rebecca's] personal information was kept confidential? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | In [William]'s town there is a large day clinic where there are several doctors and nurses. When [William] has a sensitive health problem he can see a male rather than a female doctor or nurse. | | | | | |
| Q7507 | How would you rate [William's] freedom to choose his health care provider? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | [Sarah] lives in [Kamf] where the local government has regular public meetings to discuss the community's health care needs. [Sarah] participates in identifying the community's most important health needs. Recently, they planned to build a clinic and asked people to help decide on the services and the best location. | | | | | |
| Q7508 | How would you rate the way health care involves [Sarah] in deciding what services it provides and where it provides them? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | [Elma] was asked to describe a serious health problem to the doctor in the middle of the waiting room. The doctor repeated everything to the nurse and all the other patients, many of whom knew her family, could overhear everything. | | | | | |
| Q7509 | How would you rate the way the health services ensured [Elma] could talk privately to health care providers? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7510 | How would you rate the way [Elma's] personal information was kept confidential? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | When [Penelope] had fertility problems, she had to travel from her small village to the closest town where they only had one male doctor. [Penelope] wanted to see a female doctor but she had no choice. | | | | | |
| Q7511 | How would you rate [Penelope's] freedom to choose her health care provider? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| | [Zana] attended her town's annual meeting on community health needs. They agreed to focus on improving health services at schools for the next year. After the meeting, [Zana] received a letter saying which schools would benefit, but the letter did not fully explain how the schools were selected. | | | | | |
| Q7512 | How would you rate the way health care involves [Zana] in deciding what services it provides and where it provides them? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

| | | | | | | |
|---|---|--------------|---------|-------------|--------|-------------|
| [Alouine] had his consultation in a small private room. During the consultation, a nurse occasionally walked in and listened to the conversation. Sometimes she forgot to close the door so people in the waiting room could overhear parts of their conversation. | | | | | | |
| Q7513 | How would you rate the way the health services ensured [Alouine] could talk privately to health care providers? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7514 | How would you rate the way [Alouine's] personal information was kept confidential? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| The national health service assigns all the people in [Ibrahim's] community to the local clinic. At the clinic they have a choice of 3 doctors. If they want to change clinics, they need to get permission. | | | | | | |
| Q7515 | How would you rate [Ibrahim's] freedom to choose his health care provider? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Tarek] lives near a town where the government recently built a new hospital. Before building, they didn't consult anyone in the town or surrounding villages about their health needs nor where to locate the hospital. | | | | | | |
| Q7516 | How would you rate the way health care involves [Tarek] in deciding what services it provides and where it provides them? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Ken] visited the doctor regularly. His doctor always took [Ken] to a private room before discussing his illness. The doctor was aware that [Ken] was very sensitive about his health condition and would never talk about it to anyone or in front of anyone without Ken's permission. | | | | | | |
| Q7517 | How would you rate the way the health services ensured [Ken] could talk privately to health care providers? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7518 | How would you rate the way [Ken's] personal information was kept confidential? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

[Nathan] lives in a town where there are lots of doctors and clinics. He tried one doctor but he didn't like him so he changed doctors. This was easy to do because he could go where he wanted.

| | | | | | | |
|-------|---|--------------|---------|-------------|--------|-------------|
| Q7519 | How would you rate [Nathan's] freedom to choose his health care provider? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
|-------|---|--------------|---------|-------------|--------|-------------|

Responsiveness Vignettes. Set D

I am now going to read you stories about people's experiences with health care services. I want you to think about these people's experiences as if they were your own. Once I have finished reading each story, I will ask you to rate what happened in the story as very good, good, moderate, bad or very bad.

Social Support to Patient and Autonomy

[Sally]'s husband could only visit her in hospital in the evenings because of the visiting hours and because he had to look after their children. Once he brought the children with him and the nurse, who was reading a newspaper, said she could not watch them for him. This meant [Sally] had fewer visits from her husband.

| | | | | | | |
|-------|--|--------------|---------|-------------|--------|-------------|
| Q7501 | For [Sally's] last hospital stay, how would you rate the ease of having family and friends visit her? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7502 | For [Sally's] last hospital stay, how would you rate her experience of staying in contact with the outside world when she was in hospital? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

When [Jasmina's] had treatment for infertility, the doctor gave her some pills and asked her to return in two weeks. He didn't ask her whether she wanted to know anything about her health condition nor suggest different alternatives that might have suited her better.

| | | | | | | |
|-------|---|--------------|---------|-------------|--------|-------------|
| Q7503 | How would you rate [Jasmina's] experience of getting information about other types of treatments or tests? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7504 | How would you rate [Jasmina's] experience of being involved in making decisions about her health care or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

When [Joseph] was in hospital he could have no visitors nor could he receive any presents from friends or relatives. The hospital had no telephones and he could not get any news from outside.

| | | | | | | |
|-------|--|--------------|---------|-------------|--------|-------------|
| Q7505 | For [Joseph's] last hospital stay, how would you rate the ease of having family and friends visit him? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7506 | For [Joseph's] last hospital stay, how would you rate his experience of staying in contact with the outside world when he was in hospital? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

| | | | | | | |
|---|--|--------------|---------|-------------|--------|-------------|
| [Mark] had a serious health problem. The doctor prescribed the best treatment for Mark but without telling him the implications on his quality of life or the cost. [Mark] felt powerless and was not given any information to help him to feel more in control. | | | | | | |
| Q7507 | How would you rate [Mark's] experience of getting information about other types of treatments or tests? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7508 | How would you rate [Mark's] experience of being involved in making decisions about his health care or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| When [Tamar] was in hospital for two weeks, her family visited her every few days during the evening visiting hours, but she was alone the rest of the day. There was little to read or do in the hospital. Every few days, the nurses brought her a portable radio to listen to for a few hours. | | | | | | |
| Q7509 | For [Tamar's] last hospital stay, how would you rate the ease of having family and friends visit her? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7510 | For [Tamar's] last hospital stay, how would you rate her experience of staying in contact with the outside world when she was in hospital? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Bob] had a broken arm. The doctor explained different ways of fixing it and then ordered some blood tests. [Bob] didn't know why he needed blood tests and was worried until the doctor explained what they were for. | | | | | | |
| Q7511 | How would you rate [Bob's] experience of getting information about other types of treatments or tests? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7512 | How would you rate [Bob's] experience of being involved in making decisions about his health care or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| When [Carol] was in hospital, she was allowed visitors at any time of the day. Whenever she needed to contact her family, work or friends, she could easily get a message delivered to them. | | | | | | |
| Q7513 | For [Carol's] last hospital stay, how would you rate the ease of having family and friends visit her? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7514 | For [Carol's] last hospital stay, how would you rate her experience of staying in contact with the outside world when she was in hospital? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| When [Sarah] wanted treatment for her swollen limbs, the nurses at her clinic discussed many possible treatments with her. They discussed all the pro's and con's of each treatment with her and then recommended one to her. | | | | | | |
| Q7515 | How would you rate [Sarah's] experience of getting information about other types of treatments or tests? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7516 | How would you rate [Sarah's] experience of being involved in making decisions about her health care or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |

| | | | | | | |
|---|---|--------------|---------|-------------|--------|-------------|
| [Jane] gave birth in hospital. Her husband was permitted to visit her every morning and evening. [Jane] was able to contact her family and friends once a day. | | | | | | |
| Q7517 | For [Jane's] last hospital stay, how would you rate the ease of having family and friends visit her? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7518 | For [Jane's] last hospital stay, how would you rate her experience of staying in contact with the outside world when she was in hospital? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| [Tara] is unhappy and has no energy. She gave birth to a baby girl two months ago and the doctor has told her that she may be suffering from depression. The doctor discussed her condition with her and then suggested that she could either try some medication or, if she prefers, talk to a counsellor. | | | | | | |
| Q7519 | How would you rate [Tara's] experience of getting in formation about other types of treatments or tests? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |
| Q7520 | How would you rate [Tara's] experience of being involved in making decisions about her health care or treatment? | 1. Very good | 2. Good | 3. Moderate | 4. Bad | 5. Very bad |



The responsiveness of a health system as a concept was defined by WHO and leading scientists as the health system's ability to meet the universal, legitimate expectations of its users (or clients, patients) with regard to non-medical aspects of the way they are treated and the environment (or setting) within which they are treated. This book analyses a rich set of 106 WHO household surveys on responsiveness, with approximately 258,000 respondents and 83 countries, to explore cross-country and cross-person comparability of the responsiveness concept. It also assesses the concept's application to a specific aspect of care, perinatal care, in the Netherlands. An array of analytical methods uncover the essential humanity and common expectations for quality care shared by people across the world and demonstrate the feasibility and relevance of measuring responsiveness for improving health policies and services.

DIGNITY

COMMUNICATION

CONFIDENTIALITY

PROMPT ATTENTION

AUTONOMY

QUALITY BASIC AMENITIES

ACCESS TO SOCIAL SUPPORT NETWORKS

CHOICE OF CARE PROVIDER