Tracking progress toward achieving Universal Health Coverage
Developing a methodology to estimate Effective Coverage across countries

Zitko Melo, Pedro

Awarding institution:
King's College London

Tracking progress toward achieving Universal Health Coverage. Developing a methodology to estimate Effective Coverage across countries

Pedro Zitko

Thesis submitted for the degree of Doctor of Philosophy
in Health Service and Population Research
King's College London, Institute of Psychiatry, Psychology and Neuroscience
August 2021

Abstract

The Universal Health Coverage (UHC) is a movement that promote health for all and is included among the Sustainable Development Goals. The World Health Organization (WHO), the World Bank and other institutions have proposed 'effective coverage' as an indicator for tracking progress toward UHC. Effective coverage is an indicator that combine two parameters: the coverage of a healthcare intervention and the quality of it. According to Shengelia et al.'s framework for effective coverage, quality must be measured through the health gains associated with the utilization of a healthcare service. However, none of these organizations above mentioned have properly measured effective coverage due to the lack of information about the quality of healthcare services.

In a systematic scoping review aimed to identify the use of effective coverage in the scientific literature, I found 128 studies preforming 246 assessments of healthcare interventions in 138 countries – 81% low and middle-income countries. Only one assessment included health gains into the parameter of quality. The aim of this thesis is to develop a practical procedure to estimate the effective coverage considering health gains.

Calculating health gains is related with valuating health states. There are two main approaches for valuing health states: those that look for social preferences or utilities (e.g. standard gamble, time-trade off, person -trade-off), and those that look for a direct measure of the health status or disability (e.g. paired comparison). The first approach produces 'health-utility weights', and the second approach produces 'disability weights'. Traditionally, disability weights are used for calculating disability-adjusted life years (DALY) a measure of burden of disease, while health-utility weights are used for calculating quality-adjusted life years (QALY) a measure used in cost-effectiveness analysis.

Using data from the Chilean National Health Survey 2009-2010 (Ch-NHS 2009-2010) and the Health States Description questionnaire included in that survey, I calculated a latent variable of disability. I argue that through a regression model applied to a latent variable of disability or health-states utilities, it is possible to estimate disability weights (or health-utility weights) for different health states associated with a disease, adjusting by comorbidities and other confounders.

The attributable fraction encompasses a family of epidemiological estimators that combine relative and absolute effect sizes. Attributable fractions have been used mainly for exploring the effect of risk factors on diseases. Using the attributable fraction metrics applied to a continuous outcome such us a latent variable of disability or health-state utility, I present a new way of calculating the burden of disability (or the loss of health-state utilities) associated with diseases. This methodological proposal would be more straightforward to be carried out than the standard methodological alternative (i.e. years lived with disability, a component of DALYs). Two approaches to calculate the burden of disability attributable to diseases are presented: the population average-level and the individual-level.

I also argue that the procedure to calculate the burden of disability (or loss of health-state utilities), described above, can be used to estimate effective coverage. I define effective coverage as the fraction of avoidable disability (or loss of health-state utilities) attributable to a disease, avoided by using a healthcare intervention. I also propose a definition for other related indicators: health benefit, quality, relative effective coverage (r-EC) and absolute effective coverage (a-EC). While effective coverage results from the combination of the coverage and quality, the r-EC results from the combination of the coverage and the health benefit (i.e., effectiveness). a-EC is defined as the fraction of the disability attributable to a disease in the entire population that is avoided by the healthcare intervention. This indicator is suitable to be combined with costs associated with healthcare services. The procedure to estimate these indicators is tested initially using data from the Ch-NHS 2009-2010 applied to the case of treatment for depressive disorder.

A more comprehensive appraisal of the performance of the procedure to calculate effective coverage and other indicators is also carried out using cross-sectional data from WHO study on global ageing and adult health (SAGE), Wave 1, undertaken between 2007-2010 in China, Ghana, India, Mexico, the Russian Federation, and South Africa. Three healthcare interventions were explored: treatment for depressive disorder, treatment for hypertension and treatment for osteoarthritis.

The methodological proposal for calculating effective coverage achieves estimating health gains into a parameter of quality using cross-sectional data. Among the strengths of the proposal developed in this thesis I highlight: (1) the concept of effective coverage is expanded through new indicators; (2) the procedure is straightforward to be implemented; (3) it

depends on only one source of information, which ensures consistency between parameters; and (4) it can be used indistinctly with different outcomes: disability or health-state utilities.

However, its main limitation is that the effect size attributable to the healthcare intervention is weak because the procedure proposed in this thesis is based on cross-sectional data. To improve the methodological proposal of this thesis, I highlight the following challenges for future research: (1) exploring other procedures to obtain a better proxy of the effect size of healthcare interventions using cross-sectional data (e.g. propensity score matching, instrumental variables); (2) including fatal consequences; (3) including an equity perspective in the outcome; and (4) exploring combining a-EC with the costs of healthcare interventions.

Regarding tracking progress toward UHC, I argue that a-EC would be a more adequate indicator than effective coverage. a-EC includes in a single metric the effectiveness of healthcare services, the coverage of it, the disability associated with a disease, and the prevalence of such disease. Moreover, a-EC can be added across different healthcare-services in a simpler way than with other indicators. Finally, it can be combined with the cost of the healthcare interventions, which is appropriate to inform decision makers.

Statement of Authorship

The proposal for measuring effective coverage was developed using secondary data from the following studies:

- Chilean National Health Survey 2009-2010 (Segunda Encuesta Nacional de Salud). Ministerio de Salud de Chile/ Pontificia Universidad Católica de Chile.
- The World Health Organization (WHO) Study on global AGEing and adult health (SAGE), wave 1.

These data bases were freely accessible at the moment of starting my PhD studies. I thank to all the participant of these studies, the researchers involved, and the institutions that supported these studies.

My PhD was funded by the Chilean National Agency for Research and Development (ANID): scholarship program /DOCTOARDO BECA CHILE/ Nº4721/2016.

Publications

Chapter two is based on the following publication:

- **Pedro Zitko** MD**,** Ioannis Bakolis PhD, Silia Vitoratou PhD, et al. Psychometric evaluation of the Health Sate Description questionnaire in Chile. A proposal for a latent variable approach for valuating health states. *Value in Health Regional Issues 2021 (accepted)*

The following publications are based on the methodology developed in chapters two and three:

- **Pedro Zitko** MD, MSc, Norberto Bilbeny, MD, Constanza Vargas, et al. Different Alternatives to Assess the Burden of Disease Using Attributable Fraction on a Disability Variable: The Case of Pain and Chronic Musculoskeletal Disorders in Chile. *Value in Health Regional Issues. 2021; 26(C):15–23*

- **Pedro Zitko**, Norberto Bilbeny, Carlos Balmaceda, et al. Prevalence, burden of disease, and lost in health state utilities attributable to chronic musculoskeletal disorders and pain in Chile. *BMC Public Health (2021) 21:937 doi.org/10.1186/s12889-021-10953-z*

The following abstracts were presented in international congresses. They are based on contents of chapters three, four and five:

- **Pedro Zitko**, Andrea Slachevsky, Daniel Jimenez. The prevalence, associated factors, and the burden on health state utilities for dementia in Chile. *lzheimer's Dement. 2020;16(Suppl. 10):e042466. doi.org/10.1002/alz.042466.*

- **Pedro Zitko**, Hai Nguyen. A new approach to measure effective coverage for assessing health systems performance. *Value in Health 2020, Vol 23, S1, PMU99. https://doi.org/10.1016/J.JVAL.2020.04.865*

- **Pedro Zitko**, Ricardo Araya. Effective coverage and quality of healthcare addressed to depressive disorder in China, Ghana, India, Mexico, the Russian Federation and South Africa. World Psychiatry Association Annual Congress.

Acknowledges

I would like to thank my supervisors, Professor Ricardo Araya and Dr. Ioannis Bakolis. I thank their support in all the steps of this thesis. I thank specially Professor Araya for accepting me as a PhD student.

I would also thank to Dr. Kia-Chong Chua (Centre for Implementation Science, King's College London), who selflessly was also available and genuinely interested in discussing the methodological aspect of my thesis.

I also express my gratitude to Hospital Barros Luco Trudeau and the Servicio de Salud Metropolitano Sur in Chile, places where I work. They provided the conditions to make my studies possible.

I also use the opportunity to thank Dr. Rafael Lozano of the Institute for Health Metric and Evaluation for introducing me to the passionate field of health metrics.

Finally, I would like to thank María, my beloved wife, and my daughters Josefa and Aurora. Thanks for their love and support. I also thank my parents, who backup my studies, regardless of the distance.

Table of Contents

List of figures

List of tables

Abbreviations

a-EC	Absolute effective coverage
AF	Attributable fraction
CFA	Confirmatory factor analysis
CFI	Comparative Fit index
Ch-NHS	Chilean National Health Survey
CRA	Comparative risk assessment
COPD	Chronic obstructive pulmonary disease
DALY	Disability-adjusted life years
Df	Degree of freedom
EAF	Exposed attributable fraction
EC	Effective coverage
EFA	Exploratory factor analysis
EQ5D	EuroQol five domains questionnaire
GBD	Global burden of disease
GFI	Goodness of fit index
HG	Health Gain
HSD	Health State Description
ICF-2	International Classification of Functioning, Disability, and Health
ID	Incidence rate difference
IHME	Institute of Health Metric and Evaluation
IoM	Institute of medicine
IR	incidence rate ration
NHS	National health survey
OECD	Organisation for Economic Co-operation and Development
PAF	Population attributable fraction
PID	Population incidence rate difference
PIF	Potential impact fraction
PTO	Person trade-off
QALY	Quality adjusted life years
r-EC	Relative effective coverage
RMSEA	Root mean square error of approximation
SAGE	WHO study on global ageing and adult health
SDG	Sustainable Development Goals
SEM	Structural equation model
SG	Standard gamble
SRMR	Standardised root mean square residual
TLI	Turker-Lewis index
TTO	Time trade-off
UHC	Universal Health Coverage
US	United States
VAS	Visual analogue scale
WB	World Bank
WHO	World Health Organization
WHS	Word Health Survey
YLD	Years lived with disability
YLL	Years of life lost

Chapter 1: Background

1.1. Introduction

This thesis covers a broad range of topics, from the performance of health systems up to the psychometrics, statistics and epidemiological aspects of measuring individual health status. In this chapter, I provide the background to two central subjects: universal health coverage and effective coverage. I also present the research question and the research objectives.

1.2. Universal Health Coverage

1.2.1 Background of Universal Health Coverage

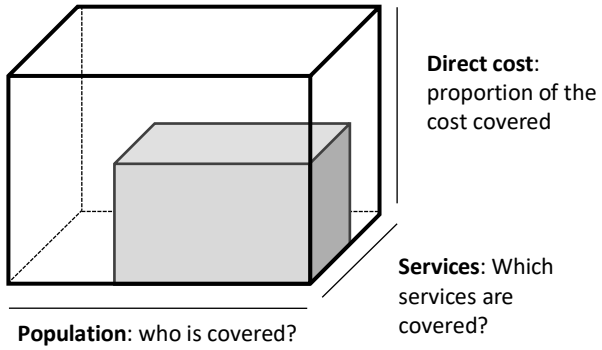
In the history of modern health systems, there have been at least three international movements that promoted universal health for all. The first ended after the second war world, with the consolidation of national health systems in a significant number of countries around the world. This occurred concurrently with the creation of the World Health Organisation (WHO) and the United Nations (UN) [1]. With the motto 'Health for all', the second corresponded to the Alma Atta declaration, in 1978, led by the WHO, which was centred on the promotion of primary healthcare as a strategy to provide for the health and wellbeing to the people, including a relevant role for communities [2, 3]. The third movement is the current universal health coverage (UHC), formally presented at the 58th World Health Assembly in 2005 with a commitment to strengthen the financing of countries' health systems[4]. The emphasis on financing can be seen in the term 'coverage', a concept initially derived from the insurance industry [3]. Unlike the 'health for all' of Alma Atta, UHC was originally concerned with how to increase the resources assigned to health care and how to avoid the catastrophic costs for people that seek health care[2, 5]. In that first instance, UHC was defined as 'everyone in the population has access to appropriate promotive, preventive, curative and rehabilitative health care when they need it and at an affordable cost' [5]. It is worth noticing in this definition the central role given to healthcare services.

In the 2008 WHO Annual Report, which was dedicated to highlighting the relevance of primary healthcare 30 years after the Alma Atta declaration, the UHC definition was changed to 'universal access to the full range of personal and non-personal health services they need, with social health protection', where social health protection is understood as 'pooling pre-paid contribution collected on the basis of ability to pay, and using these funds to ensure that services are available, accessible and produce quality care for those who need them, without exposing them to the risk of catastrophic expenditure'[6]. Interestingly, this new definition mentions 'non-personal health services', and introduces the concept of quality.

In that report, WHO proposed three technical dimensions that should be considered to move countries toward UHC: the proportion of people covered (insured), the range of services (benefits) covered and the proportion of the total cost that comes from a pooled fund (public health expenditure) [6] (see Figure 1). This three-dimensional concept of UHC was

relaunched in the 2010's WHO report, addressed to health system financing [7]. In this last report, UHC was defined simply as when 'all people have access to services and do not suffer financial hardship paying for them' [7].

Figure 1.1 Three dimensions to move toward UHC



Adapted from '*The World Health Report: health systems financing: the path to universal coverage'*. 2010, World Health Organization. The grey cube represents an example of the level of universal health coverage in a country.

Despite the political usefulness of a broad concept for UHC, several criticisms have arisen. Among them, it highlights the ambiguity of its definition and the lack of a consistent conceptual framework that allows creating a robust indicator able to monitor any progress [8]. Additionally, it is said that UHC neglects the role of the social determinants of health and the role of the extra-sectorial actors, which results in removing attention about issues related to health inequalities [2, 8]. The concept of coverage could also be understood as the mere accessibility to healthcare services since it does not include their use, opportunity and quality of them, at least not explicitly [8]. Finally, although UHC is closely related to the financing of health systems, its definition does not provide a clear proposal about how to structure health services to fulfil the provision of care [2], or a clear plan about how to strengthen health systems as a whole [2, 8].

In 2012, the action of governments toward UHC was encouraged through a Resolution of the 67th United Nations General Assembly, which called for 'accelerating the transition towards universal access to affordable and quality health care services' [9]. More recently, UHC was included as an indicator in the Sustainable Development Goals (SDGs) launched in 2015 [10].

1.2.2 UHC in SDG

Into the SDG's third goal ('Ensure healthy lives and promote wellbeing for all at all ages'), UHC indicator (3.8) was defined as: 'Achieve universal health coverage, including financial risk protection, access to quality essential healthcare services and access to safe, effective, quality and affordable essential medicines and vaccines for all´ [10]. In this definition, the emphasis is once again on healthcare provision and even more limited to essential medicines and vaccines. However, the concept of quality is retained as a desirable attribute of health care interventions.

Finally, a new operational definition for UHC was provided in 2017 in the formal presentation of SDG indicators. In this way, indicator 3.8 was split into two sub-indicators: one addressed tracking the coverage of health services and another was concerned with financial protection (see Table 1.1).

Table 1.1. Indicator 3.8 of Sustainable Development Goals according to the Resolution adopted by the United Nations General Assembly on 6 July 2017[11]

Indicator	Sub-indicator
3.8. Achieve UHC, including financial risk protection, access to quality essential health-care services and access to safe, effective, quality and affordable essential medicines and vaccines for all.	**3.8.1.** Coverage of essential health services (defined as the average coverage of essential services based on tracer interventions that include reproductive, maternal, new-born and child health, infectious diseases, non-communicable diseases and service capacity and access, among the general and the most disadvantaged population).
	3.8.2. Proportion of population with large household expenditures on health as a share of total household expenditure or income.

In this operational definition, the health care interventions were again broadened in terms of the diversity of conditions (infectious, reproductive, maternal, new-born and child, and non-communicable diseases), although limited to a set of traceable health needs. Additionally, under the concept of 'coverage of essential health services', it was introduced the notion of capacity of services and a specification about the distribution of the coverage in the population.

The second sub-indicator is used to identify catastrophic and impoverishing expenditures, defined as the spending on health care that can lead to the household members to suffer a significant burden or move them under the poverty line, respectively [6]. However, for this version of the indicator, no specific threshold is given for the fraction of the total household income that if spent can lead to a financial catastrophe.

The use of two sub-indicators separates the dimensions presented in Figure 1, leaving coverage and the range of interventions to the indicator 3.8.1, and the financial dimension to the indicator 3.8.2. These indicators are chosen as a technical definition for UHC.

1.2.3 How UHC is monitored

Since the SDGs were launched, several organisations have proposed ways to monitor them. For indicators related to health, two major institutions have proposed specific measurements. One of these is the WHO, which, in collaboration with the World Bank (WB) [12] has planned specific measures for indicators 3.8.1 and 3.8.2 of UHC [13, 14].

In relation to the first sub-indicator about coverage of essential health services, WHO and WB have declared that the ideal indicator is effective coverage, also called effective service coverage. This indicator is understood as 'the proportion of people in need of services who receive services of sufficient quality to obtain potential health gain' [15, 16]. They provide an example of treatment for HIV, pointing out that it is not only important knowing the proportion of infected people that are receiving antiretroviral therapy (i.e., treatment coverage), but also the fraction that has achieved viral suppression (i.e., treatment effective coverage).

However, they discarded the use of effective coverage for all selected conditions when tracking progress toward the UHC, mainly due to the difficulty in obtaining metrics of quality for each kind of healthcare intervention, especially when required for a high number of countries [15, 16]. Table 1.2 presents the composite index proposed by WHO and WB, called 'UHC service coverage'. This index is made up of twelve sub-indicators. Of these, only two correspond to the measurement of effective coverage, while seven match the concept of service coverage and the remaining seven are only a proxy of service coverage. Family planning, which is one of those using effective coverage, measured quality as the proportion

of women who have acceded to a 'modern method', over those who accessed any method of family planning [16]. Information about family planning was extracted from national surveys. In the case of tuberculosis effective treatment coverage, the measurement was calculated through administrative records that combined an inference of incidence, the number of people detected and the number of individuals who successfully completed the treatment [16].

Table 1.2. UHC service coverage index according to the World Health Organization and World Bank.

Reproductive, maternal, newborn and child health (RMNCH) 1. Family planning (FP) ESC 2. Antenatal care4 + visits (ANC) SC 3. Child immunisation (DPT3) SC 4. Care seeking suspected pneumonia (Pn) SC	RMNCH = $1/(FP \bullet ANC \bullet DTP3 \bullet Pn)^{1/4}$	UHC service coverage index =$(RMNCH \bullet Infec \bullet NCD \bullet SCA)^{1/4}$
Infectious disease control (IDC)* 1. TB effective treatment (TB) ESC 2. HIV treatment (ART) SC 3. Insecticide-treated nets (ITN) SC 4. At least basic sanitation (WASH) SC	IDC = $1/(TB \bullet ART \bullet ITN \bullet WASH)^{1/4}$	
Non-communicable diseases (NCD)† 1. Normal blood pressure (BP) Proxy 2. Means fasting plasma glucose (FPG) Proxy 3. Cervical cancer screening (CeCa) SC 4. Tobacco non-smoking (To) Proxy	NCD = $1/(BP \bullet FPG \bullet To)^{1/3}$	
Service capacity and access (SCA)†† 1. Hospital bed density (Hosp) Proxy 2. Health worker density (HWD) Proxy 3. Access to essential medicines Proxy 4. IHR core capacity index (IHR) Proxy	SCA = $1/(Hosp \bullet HWD \bullet IHR)^{1/3}$	

Adapted from [16]

ESC*: effective service coverage/* SC*: service coverage/* IHR*: international health regulations.*
TB*: tuberculosis.*
** In the case of countries with a low risk of malaria, the calculi is performed: Infec = 1/(*TB ● ART ● WASH*)$^{1/3}$*
***†** Cervical cancer screening is omitted in the calculi because of lack of information/* BP *and* FPG *are rescaled to a scale between 0 and 1.*
***††** Access to essential medicines is omitted in the calculi because of lack of information. Continuous values of hospital beds,* HWD *(health workers density) and* IHR *(international health regulations) are calculated using threshold values and rescaled between 0 and 1.*

All sub-indicators are weighted equally. A final index is obtained using a geometric mean of all indicators. The authors recognise that using a geometric mean on equally-weighted sub-indicators is not robust but point out that a single index is easier to be communicated to decision-makers [13]. The final indicator includes different metrics, including the density of health workers, mean of FPG and HIV treatment coverage.

Regarding indicator 3.8.2, about financial protection, WHO and WB calculated the proportion of people who annually incurred in an out-of-pocket payment for any health care intervention larger than 10% and 25% of the total consumption of their household [16]. This indicator, presented in their second SDG monitoring report (2017), was a modification of that introduced in their first report in 2015, although conceptually, the measurement remained unchanged. The authors point out that there is a wide consensus on how to measure financial protection [13].

WHO and WB are not the only institutions interested in measuring UHC. Other organisations that have tracked the health-related SGDs include the Institute for Health Metrics and Evaluation (IHME) dependent of the University of Washington and financed by the Bill & Melinda Gates Foundation [http://www.healthdata.org/about/history, consulted in July 2020]. This institution, created in 2010, has a mission to 'improve the health of the world's populations by providing the best information on population health' [http://www.healthdata.org/about, consulted in April 2020]. It is known for its global burden of disease reports which, in collaboration with *The Lancet*, are published periodically [17]. The IHME was created by professionals that started this work at WHO where they founded the Global Burden of Disease Project in the early 2000s, and other relevant studies such as, for example, the first worldwide health service performance assessment [18]. Its leader, J.L. Christopher Murray, is also known for his authorship of the WB's first study of the burden of disease published in 1993, where the metric of Disability-Adjusted Life Years (DALY) was presented [19].

The way that IHME is tracking progress toward UHC is presented in Table 1.3.

Table 1.3. UHC index according to the Institute of Health Metric and Evaluation [20]

Definition	**Further details**
Coverage of essential health services, as defined by a UHC index of the coverage of nine tracer interventions and risk-standardised death rates from 32 causes amenable to personal health care	Tracer interventions included: • vaccination coverage (coverage of three doses of diphtheria-pertussis-tetanus, measles vaccine and three doses of the oral polio vaccine or inactivated polio vaccine); • met need for modern contraception; • antenatal care coverage (one or more visits and four or more visits); • skilled birth attendance coverage; • in-facility delivery rates; and

	• coverage of antiretroviral therapy among people living with HIV. The 32 causes amenable to personal health care, which compose the HAQ Index, included: • tuberculosis, diarrhoeal diseases, lower respiratory infections, upper respiratory infections, diphtheria, whooping cough, tetanus, measles, maternal disorders, neonatal disorders, colon and rectum cancer, non-melanoma cancer, breast cancer, cervical cancer, uterine cancer, testicular cancer, Hodgkin's lymphoma, leukaemia, rheumatic heart disease, ischaemic heart disease, cerebrovascular disease, hypertensive heart disease, peptic ulcer disease, appendicitis, hernia, gallbladder and biliary diseases, epilepsy, diabetes, chronic kidney disease, congenital heart anomalies and adverse effects of medical treatment. These 41 individual inputs are presented on a scale from 0–100, with 0 reflecting the worst levels observed between 1990 and 2016 and 100 reflecting the best observed during this time. The arithmetic mean of these 41 scaled indicators is used, to collectively capture a wide range of essential health services on reproductive, maternal, newborn and child health; infectious diseases; NCDs; and service capacity and access.

Adapted from [20]

The index proposed by the IHME combines nine indicators of coverage (coverage of three vaccines, family planning, antenatal care, skilled birth attendance, delivery in healthcare facility rate and antiretroviral therapy coverage) and 32 mortality rates for amenable causes [21, 22]. All sub-indicators of coverage and mortality are standardised on a scale from 0 and 100, according to maximums and minimums valued observed since 1990. Again, the arithmetic mean is used to consolidate the index to the national level, providing the same weight to each sub-indicator.

The IHME has declared that the ideal index to track the UHC should be effective coverage [23]. However, as is seen in Table 1.3, its proposal for monitoring UHC is far from that concept. The concept of effective coverage is developed later in this chapter.

In summary, UHC is an idea with a worldwide movement supporting it to enhance access to quality health services, protecting people from financial risks. Major institutions including WHO, WB and IHME have proposed effective coverage as an ideal indicator to measure

progress in that direction. However, the UHC's indices that they have suggested are still far from the concept of effective coverage.

1.2.4 Effective coverage out of the context of UHC

In 2000, WHO published the results of the first effort to assess the performance of health systems worldwide [18]. Despite this huge milestone the report was criticised for its methodological approach [24, 25]. As a result, in 2002 WHO requested an external evaluation [26] and a new plan to refine and improve the framework for assessing the performance of health systems [27]. In both instances, 'effective coverage of interventions' emerged as a desirable metric for measuring the performance of the health systems [28].

In a more recent report (2013), the European Observatory of Health Systems and Policies also proposed to use this metric for comparing health systems across countries. However, they recognised the difficulties inherent in its implementation, especially the challenge to quantify the quality of the healthcare interventions [29].

In the next section, I will present the definition of effective coverage, referencing the two main works into which this metric has been conceptually developed.

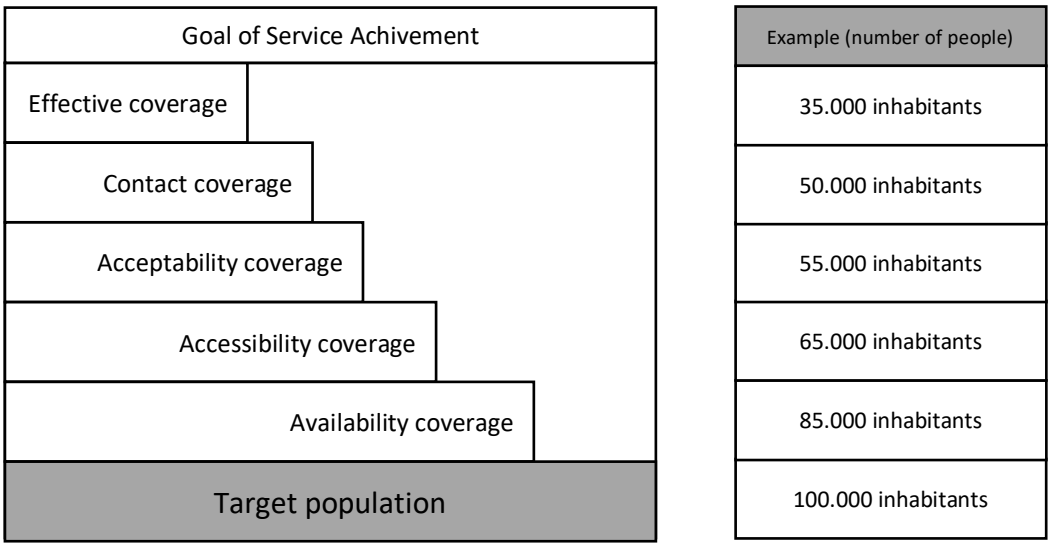
1.3 The concept of effective coverage

1.3.1 First definition: T. Tanahashi

The concept of effective coverage was firstly proposed by T. Tanahashi in 1978, a scientist from the Division of Strengthening of Health Services at WHO Tanahashi defined effective coverage as 'the fraction of a target population that receives a satisfactory or effective health intervention for [certain] condition' [30]. In that definition, Tanahashi highlights its difference from the common concept of contact coverage; the fraction of a target population that makes contact with a provider that delivers a certain health intervention [30].

For Tanahashi, the target population is defined as those that can benefit from the healthcare service under assessment, which is in agreement with other authors [31]. However, contact and effective coverage are only the final steps in the process of assessing coverage from the perspective of health services. Other prior types of contact identified are: 'available coverage' (i.e., the amount of service that can be made available to the target population); 'accessibility coverage' (i.e., the number of people who can reach and use the service); and 'acceptability coverage' (i.e., the number of people who are willing to use the accessible service) (see Figure 1.2).

Figure 1.2. Coverage diagram according to T. Tanahashi, 1978 [30]



Adapted from Tanahashi T., Bulletin of the World Health Organization 1978;56(2):295-303

Although Tanahashi does not mention it, from his article and the diagram that he presents, we can see each step of coverage as a conditional fraction from the previous step. That means that accessibility coverage is a fraction of the target population that has availability coverage, acceptability coverages is a fraction of the target population that has availability and accessibility coverage and so on.

Another important element is that Tanahashi makes a difference between the actual coverage, calculated using the people that have already received the health intervention and the potential coverage, calculated considering the people to whom the health service can be provided. While the actual coverage depicts the output of the health service, the potential coverage describes its capacity. According to this framework, utilization is the ratio between the output and capacity. Moreover, from the five steps of coverage, availability, accessibility and acceptability would be part of the potential coverage (capacity), and contact and effective coverage would be part of the actual coverage (output). Using the numbers given as an example in Figure 2, the potential coverage can be calculated, dividing the number of people that have availability, accessibility and acceptability by the size of the target population (i.e., 55.000/100.000 = 55%). In contrast, the actual coverage, using as a reference the contact coverage, is 50.000/100.000 = 50%. Utilization, according to Tanahashi, would be 0.50/0.55, or 90.1%. The difference between coverage and utilization is significant; while the former uses the target population as a reference, the concept of utilization uses the capacity of the health service as a reference. Tanahashi also describes other kinds of coverages that can be calculated using his conceptual framework, either using a denominator from different to the target population (provision-specific coverage) or using a subgroup of the target population.

However, above all the complexities that Tanahashi added with his conceptual framework, two central elements are important in this thesis. The first one, is that when Tanahashi defines effective coverage through a 'satisfactory or effective health intervention', he is referring to a binary outcome: the target population obtains or does not obtain a satisfactory or effective health intervention. The assumption of the effectiveness indicator as a dichotomous variable has significant metric consequences, which will be seen later when I analyse how effective coverage has been measured in the scientific literature.

The second central element is, as Tanahashi says, that using the number of people from the target population as a denominator is a metric simplification. Depending on the intervention,

it may be more convenient to use the expected 'number of cases' (e.g., if the intervention can be provided more than once to each individual), or even conceptually better, the 'amount of need'. The concept of the amount of need, according to Tanahashi, can be called also the 'target service', which does not depend on the size of the target population, but on the total potential benefit that a health intervention can provide at the population level. Opting for this approach has metric consequences which will be discussed later.

1.3.2 Second definition: B. Shengelia et al.

After the publication of Tanahashi's work, few other articles used the concept of effective coverage until 2005 when B. Shengelia et al. published an update of the conceptual framework of effective coverage, which also sought to integrate other related concepts such as access, utilization, and quality of the healthcare [32]. Shengelia at the time when he published his proposal, was part of the staff of the WHO's Department of Chronic Diseases and Health Promotion, and other co-authors belonged to other institutions such as the Global Health Institute at the Harvard University. Several of these authors, later migrated to the IHME, including Christopher Murray.

Shengelia et al. defined effective coverage as: 'the fraction of maximum possible health gains an individual with a healthcare need can expect to receive from the health system'. In contrast with the Tanahashi's definition, the concept of 'target population' as the denominator of coverage is no longer used, rather the 'health gain', which is closer to the idea of 'amount of need', is adopted. The idea of 'satisfactory or effective intervention' from Tanahashi is replaced by the concept of maximum possible health gain, which is now circumscribed up to the level of a health system. Finally, instead of using the population level, the definition is now given at the individual level.

Shengelia et al., formalised the definition of effective coverage using the following equation:

$$EC_{ij} = Q_{ij}\, U_{ij} \mid N_{ij}=1 \qquad \text{[Eq1.1]}$$

where EC_{ij} is the expected effective coverage for the individual *i* and the healthcare intervention *j* and Q_{ij} is the expected quality of the healthcare intervention *j* for the person *i*. $U_{ij} \mid N_{ij}=1$ means that the utilization of the healthcare intervention *j* for the individual *i*, is

conditioned to the true need of such intervention, understanding the true need as the potential to benefit from it [31].

Note that $U_{ij} \mid N_{ij}=1$ is equivalent to the concept of contact coverage from Tanahashi's framework. Hence, effective coverage for Shengelia et al. could be also understood as the product of contact coverage and the quality of the healthcare intervention. Consequently, the concept of utilization for Shengelia et al. differs from that proposed by Tanahashi, especially because for Shengelia et al., utilization is no longer linked with capacity, a previously discussed concept (see section 1.3.1). Although, conditioning utilization to the true need is not exclusive of Shengelia et al.'s framework, it is explicitly present in the equations that they propose.

Conditioning the utilization of a healthcare intervention to the true need is important since, frequently, when contact coverage is estimated, the source of information (or databases) differs for the people with a need and people with utilization. Thus, using different sources of information (or databases) precludes ensuring that the numerator of the indicator (people receiving the health service) is included in the denominator (people in need)[33]. For example, if we wanted to estimate the coverage of the treatment for a disease such as depression, health systems can calculate the number of prevalent cases using national surveys or results from a literature review (for instance: prevalence of depression = 9%). That number can be applied to a reference population to calculate the target population (e.g., 9% * 13.5 millions of inhabitants = 1,215,000 people with normative need of treatment for depression). Later, the number of people admitted to the national mental health programme can be extracted from administrative records (i.e., 335,000 individuals) [34]. Using this approach, coverages over 100% can be found because the source of information, the time frame and the criteria for identifying cases can vary. In addition, people without a real need (e.g., without depression) could be counted as receiving the treatment [23].

For Shengelia et al., the concept of quality is essential, and it is defined as:

$$Q_{ij=} \quad \frac{\sum_{k=1}^{n} HG_{ijk} U_{ijk}}{\sum_{k=1}^{n} U_{ijk} HG_{ijk} \mid P_k = P_k^{max}} \qquad \text{[Eq2]}$$

where HG_{ijk} is the health gain for the individual *i*, healthcare intervention *j* and the healthcare provider *k*. The expression in the denominator $HG_{ijk} \mid P_k=P_k^{max}$, means the potential health gain for the individual *i*, healthcare intervention *j* and the healthcare provider *k,* assuming that all providers are delivering the intervention optimally. U_{ijk} is a set of probabilities of choosing among *n* different providers available for each individual *i* and healthcare intervention *j*. It is worth noticing that Q_{ij} is just a proportion between the actual health gains across different providers and the optimum health gain, resulting in a number between 0 and 1.

The Shengelia et al. framework allows us to aggregate the effective coverage calculated for each individual and intervention and then aggregate the effective coverage for different interventions in a single measure of total effective coverage, theoretically for the whole system. The only precaution that must be considered is weighting each effective coverage by the magnitude of the health gain offered by each intervention to each individual. Shengelia et al., as an example of health gain measurement, use the metric of healthy life years [35], but they mention that health gains can also be extracted from similar sources used in the cost-effectiveness analysis (e.g., DALY, quality-adjusted life years). The topic of health gain will be addressed exhaustively in the following chapters.

Another relevant element is that, according to Shengelia et al., the quality of healthcare interventions depends on two main factors: the implementation and choice of the adequate intervention and the adherence by people who receive the service. This is a specific understanding of the concept of quality. How much it is related to other definitions of quality will be discussed in a subsequent chapter. It is important to notice that the Tahanashi's approach also allows estimations of coverage adjusted by adherence or other criteria of quality.

In summary, there are two main seminal pieces of work that have shaped our current theoretical frameworks about effective coverage. The main difference between them is how the 'effective' component is understood in each model. While Tanahashi seems to assume a dichotomous construct, Shengelia et al. provide a more comprehensive approach including the idea of a potential health gain measured at the level of providers of a healthcare intervention. In that way, Shengelia et al. are also proposing a metric for the concept of

quality of healthcare. These elements make Shengelia et al.'s approach appealing as a starting point.

In the following section, I will present a literature review that shows how effective coverage has been used in the scientific literature.

1.4. The usage of effective coverage in the scientific literature: a scoping review

To assess how effective coverage has been used in the scientific literature, I carried out a scoping review to:

1. Identify health interventions for which effective coverage has been measured;
2. Identify the context (year and country) where effective coverage has been measured; and
3. Recognise the gap between the empirical measurement of effective coverage and Shengelia et al.'s conceptual framework, especially in the concepts of utilization and quality.

These objectives can be summarised in the following research question: *On what interventions, in which context and to what extent does the measurement of effective coverage in the existing literature match Shengelia et al.'s conceptual framework?*

A scoping review aims to map the relevant literature about a topic area, but not as accurately and specifically defined as in a systematic review. Because a scoping review is interested in a general appraisal, it has less restrictive inclusion criteria and commonly does not assess the quality of the selected studies [36]. Since I wanted to obtain a general evaluation of how effective coverage is used in the scientific literature, I considered that a scoping review was adequate to this aim.

Because utilization and quality are the critical elements of Tanahashi and Shengelia et al.'s conceptual frameworks, I have paid special attention to how the studies have dealt with these definitions. This is relevant in the case of the quality parameter, since its measurement has been described as the main limitation for more extensive use of the effective coverage indicator [16, 29].

1.4.1 Methods of the scoping review

The search for articles was carried out through three databases: PubMed (from 1946), EMBASE (from 1974) and Global Health (from 1973) and was run off on the 17th of April 2020,

using the OVID platform. Essentially, I looked for articles in whose title, abstract or keywords the term 'effective coverage' was included (see Table S1.1 from supplementary material for the electronic search strategy). The search was complemented with references from other systematic or scoping reviews identified on the topic and from special journal collections devoted to the measurement of UHC [37].

I considered only two inclusion criteria: 1) measurement of effective coverage, at least for one of the assessed interventions, and 2) studies selected by other systematic reviews aimed to identifying effective coverage studies. There was no restriction in terms of date, language, study design or type of record (article or abstract of a congress). Studies, where effective coverage was approached with models different from Tanahashi or Shengelia et al., were excluded.

The report adheres to the criteria suggested by the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) [38]. A protocol for this scoping review was presented in a previous document in the Upgrade of this thesis. The protocol was not registered on any platform. The analysis and selection of records and articles were carried out by only one researcher (PZ). No contact with any author was made to collect additional information from the selected articles. The design of the spreadsheet used to extract the information from articles is presented in the supplementary Table S1.2. I collected the year of publication, the type of intervention analysed and the country where the assessments were carried out. Additionally, I identified the definition of utilisation used and I checked if it fulfilled the criteria that the number of people who received the intervention was included in the denominator of people in need for such intervention. I also registered the way that quality, understood as a health gain, was measured and then classified in one out of six categories. For practical reasons, I used as a reference the International Classification of Functioning, Disability and Health (ICF-2) from WHO [39]. In this classification, the functioning and disability can be split into three constructs: 1) functions and body structures (e.g. a person has a joint deformity and has a reduced range of joint movement); 2) activities and participation (e.g. a person can walk 100 metres on a flat surface or around where they live); and 3) a consolidate measurement of the health status or disability from the capacity or performance to achieve activities or participation on several domains (e.g. a score between 0 and 1, where 0 means healthy and 1 death) [40]. This classification and the underpinning concepts are developed in chapter two. In addition to these categories, three others were included: content of health care, a health facility index

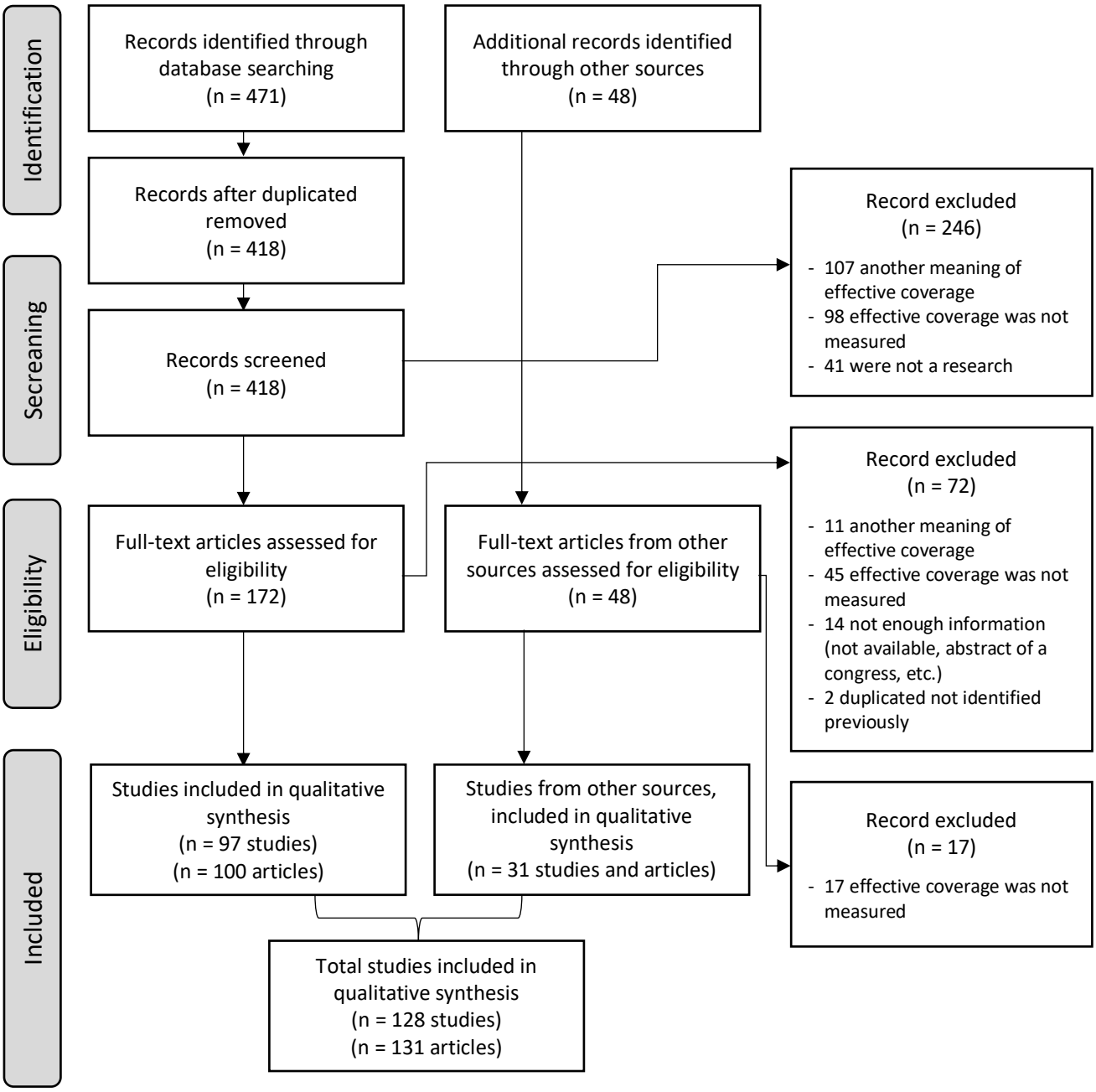
of quality and the incidence of new diseases or health conditions. These categories can also be included in the ICF-2 framework under the concepts of morbidities and contextual factors (see chapter two). Finally, I also explored if the studies included all types of coverage proposed by Tanahashi: availability, accessibility and acceptability, apart from contact coverage and effective coverage, already registered as utilisation and quality (i.e., health gain). Affordability also was added because Shengelia et al. included this element of coverage as a distinctive step and different from accessibility and acceptability.

1.4.2 Results of the scoping review

In Figure 1.3, the flow diagram for the selection of studies is shown. The search identified 471 records, corresponding to 418 unique records after removing duplicated. Abstracts were screened and 107 were eliminated since the effective coverage was used with other meanings: social media, electromagnetic fields and the coverage spectrum of antibiotics, among others. Additional studies were eliminated because they were not research but narrative reviews, while others used the term effective coverage only to contextualise their research.

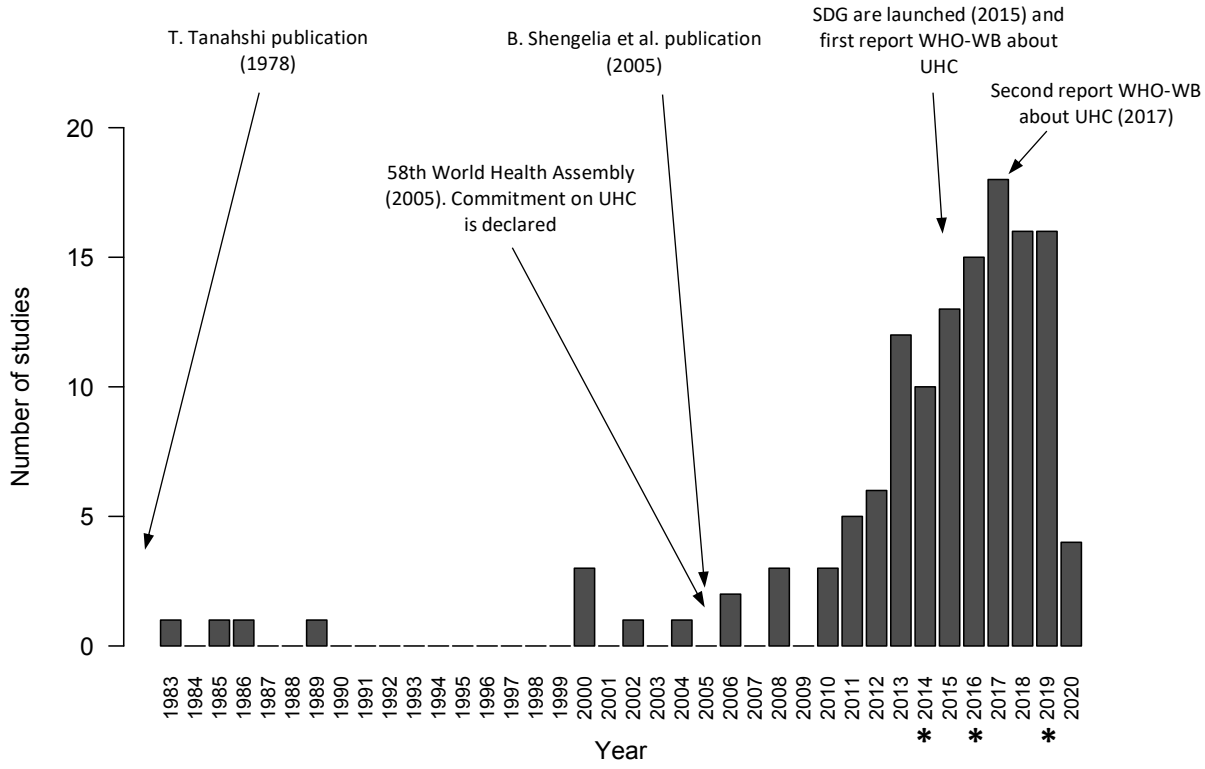
Some 172 full-text articles were reviewed and 72 were discarded for similar reasons. The review identified two systematic and one scoping review related to effective coverage. One of the systematic reviews from 2014 sought to estimate the coverage of mental health programmes [41] and contributed seven more studies to the review. The second systematic review [42], published in 2019, addressed the coverage of reproductive, maternal, newborn and child health and nutrition interventions. It included 36 studies, 19 of which were not identified by the search. Finally, a scoping review from 2016, sought to explore the use of effective coverage in the assessment of the health system's performance. It selected 18 studies, from which our review identified 15. The total number of studies included in the analysis was 128, equivalent to 131 articles. Three studies had more than one publication.

Figure 1.3 Flow diagram of the scoping review on effective coverage



The number of studies selected by the year is presented in Figure 1.4. In that figure, also I point out some relevant milestones related to UHC and the development of the metric of effective coverage. It can be observed that the studies that measure effective coverage are relatively recent, mainly from the last seven years.

Figure 1.4 Number of studies selected by year of publication.



UHC: universal health coverage/ SDG: *Sustainable Development Goals/* WHO*: World Health Organisation/ WB: World Bank. / * points out the year where a systematic or scope review about effective coverage was published.*

Considering each study can involve the assessment of more than one intervention in more than one country, in total 246 effective coverage assessments were analysed, which accounted for 408 country-specific assessments. Some 46 different interventions were identified and carried out in 138 countries, of which 26 (18.8%) were a high-income countries according to the World Bank's 2019 classification. In terms of country-specific assessments, only 38 (9.4%) were in a high-income country.

The 46 interventions identified were classified into seven groups according to the disease that they addressed: infectious diseases; maternal, neonatal conditions and general child-care; nutrition; non-communicable diseases (excluding mental health interventions); mental health; injuries and violence; and sanitation and other preventive interventions.

Table 1.4 presents a summary of the findings. Maternal, neonatal and general child-care interventions accounted for more than a third of all assessments of effective coverage (87/246 = 35.4%). They involved 96 countries, with only one from a high-income country. The interventions included by sub-type were family planning, antenatal care (overall and 18

specific activities), obstetric care, skilled birth attendance, post-natal care, preventive health control in the child, child health services and maternal and child health services (see supplementary Table S1.3-4 for details). In 77% of all assessments, utilisation was conditioned to people in need of the intervention, while in the remaining, the number of people was estimated indirectly. The indicator of health gains (quality) was determined mainly through contents of care, for instance: pregnant women had at least four antenatal visits to the healthcare provider, or the delivery was attended in a hospital. Also, a quality index based on attributes observed or registered from health facilities was used (see Table S1.5 for a breakdown for each sub-type). Overall, the assessment of the different types of coverage according to Tanahashi categories was low. It was more common for maternal, neonatal and general child-care interventions, especially in availability (23.0% of the assessments) and accessibility (16.1%).

Interventions addressing infectious diseases accounted for little less than a third of all effective coverage assessments (i.e., 71 assessments, 38 studies), encompassing: immunisation; treatment for acute respiratory infections in children; treatment for diarrhoea in children; massive drug administration for lymphatic filariasis; vector control, diagnosis and treatment for malaria; counselling, testing and treatment for HIV; screening and treatment for tuberculosis; and massive drug administration for trachoma. Less than 4% of assessments were conducted in a high-income country. Utilisation was measured in agreement with Shengelia's criteria in 90% of the cases. In almost a quarter of the assessments, it was not possible to assess quality and the authors reported only contact coverage. When the quality measure was available, it was used as a proxy for content of care or a quality index of the health care facility. Only in 11.3% of cases were biomarkers or a functionality measure used, such as seroprevalence (e.g., immunisation) or clinical-parasitological cure (e.g., malaria treatment). In a third of the assessments, acceptability coverage was described. More detailed information is available in supplementary Tables S1.6-8.

Table 1.4. Summary statistics (frequency and percentages) of selected studies that measured effective coverage, according to the group of diseases that the intervention addressed.

	Infectious diseases	Maternal, neonatal and general child-care	Nutrition	Non – communicable diseases	Mental health	Injuries and violence	Sanitation & others
N assessments (total =246)	**71**	**87**	**18**	**48**	**17**	**1**	**4**
% of total assessments	**28.9**	**35.4**	**7.3**	**19.5**	**6.9**	**0.4**	**1.6**
Number of studies	38	41	15	18	10	1	2
Number of countries involved	51	96	8	64	8	1	1
% of the countries are high-income	3.9	1.0	0.0	32.8	62.5	0.0	0.0
Utilisation is conditioned to a normative need (%)*	90.1	77.0	100.0	83.3	52.9	100.0	100.0
Type of measurement for quality*							
Quality (health care provider) (%)	12.7	29.9	0.0	0.0	0.0	0.0	0.0
Quality (incidence of a disease) (%)	2.8	8.0	0.0	6.3	0.0	0.0	0.0
Quality (content care) (%)	50.7	54.0	50.0	18.8	5.9	100.0	25.0
Quality (biomarker and functionality) (%)	11.3	1.1	50.0	56.3	41.2	0.0	0.0
Quality (activity/ participation) (%)	0.0	0.0	0.0	2.1	11.8	0.0	0.0
Quality (health status/disability) (%)	0.0	0.0	0.0	0.0	5.9	0.0	0.0
Type of coverage studied according to Tanahashi *							
Availability (%)	14.1	23.0	0.0	2.1	0.0	0.0	0.0
Accessibility (%)	7.0	16.1	0.0	4.2	0.0	0.0	0.0
Affordability (%)	4.2	1.1	0.0	0.0	0.0	0.0	0.0
Acceptability (%)	29.6	4.6	5.6	6.3	0.0	0.0	0.0
Contact (%)	97.2	98.9	100.0	100.0	100.0	100.0	100.0
Quality (%)	74.6	93.1	100.0	83.3	64.7	100.0	25.0

** the percentages are based on the number of assessments.*

The effective coverage for non-communicable diseases was included in nearly 20% of all assessments (48 assessments, 18 studies) and a third were carried out in a high-income country. Utilisation was measured ensuring that people with needs were included in the denominator in more than 80% of cases. The interventions included were: treatment and preventive activities for diabetes; screening and treatment for hypertension; treatment for hyperlipidaemia; screening for cervical and breast cancers; and treatment for vision disorders. Some studies also included treatment for angina, asthma and arthritis, but authors provided a measure of quality for none of them, only reporting contact coverage. Most of the studies explored the effective coverage for the treatment of hypertension and diabetes, using as a health gain measure the fraction of people that reach levels of blood pressure or glycaemia under recommendations (i.e., biomarkers of functionality). Only one study addressed visual disorders [43], exploring the health gain as activities/participation according

to the ICF-2 criteria: 'report no near or far visual impairment when wearing glasses or contact lenses'. More detail is available in supplementary Tables S1.9-11.

Mental health interventions accounted for only 6.9% of all assessments (10 studies) but these were more common among the high-income countries participating (62.5%). They were: treatment for depression, drug and alcohol use disorders, obsessive-compulsive disorder, psychosis and severe mental disorders, epilepsy and the mental health programme as a whole. This category was the one with the lowest rate of utilisation following the criterium of Shengelia et al. The number of people in need was commonly estimated through surveys, while the contact with health services was extracted from administrative records. However, it was the only type of intervention where at least one study calculated health gains using a health status or disability measurement, DALY [44]. Other studies used recovery [45], a close concept to activities and participation, and several studies used symptomatology criteria to measure health improvements (see supplementary Tables S1.12-14 for more information).

The interventions to prevent or treat nutritional conditions included 18 assessments in eight different countries, none of them a high-income country. In all studies, utilisation considered only people with need of the intervention, generally defined by the age (children) and sex (women) of the target population. The quality was measured usually through the content of care (e.g., the children consume the supplement adequately) or through a biomarker of functionality such as achieving the required intake according to recommendations (see supplementary Tables S1.15-17).

There was only one study that explored the effective coverage for an intervention addressed to injuries and the health consequences of violence, which was directed to domestic violence against women. The quality was measured as 'the staff recommended reporting the perpetrator to the police authorities following country regulations' [46].

Finally, four assessments, three of which were carried out in China [47], explored access to and the use of sanitary toilets, access to safe drinking water and a smoking cessation intervention. Only one of these assessments evaluated quality, understood as the compliance for sharing toilets facilities[48] (see supplementary Tables S1.21-23).

A complete description of each one of the 131 selected articles is presented in the supplementary Table S1.24.

1.4.3 Discussion about results of the scoping review

1.4.3.1 Main findings

Effective coverage is a metric recently applied in the scientific literature. It has been used mainly to evaluate maternal and child, infectious and nutritional interventions (71.6% of all assessments) and it is applied principally in low- and middle-income countries. Although, the measurement of quality was included in 84.1% (207/246) of all assessments, in more than a half (139/207 = 67.1%) it was quantified through contents of care or a quality index of the healthcare facility, which are far from the original concept of health gain proposed by Shengelia et al. Only one study measured the concept of health gain based on the Shengelia et al. definition, while three more used close constructs such as activities and participation. The complete sequence of different steps of coverage proposed by Tanahashi has been little studied.

1.4.3.2 Meaning of findings

Although, non-communicable diseases account for more than 60% of the global burden of diseases [49], in our review, the assessments of effective coverage of healthcare interventions addressing that group of diseases (including mental health interventions) added up to only 26.4%. On the contrary, although almost three-quarters of the assessments of effective coverage were about a maternal and child, infectious and nutritional interventions, the disease burden of these conditions accounts for less than 30% of the total global burden of diseases [49]. There is thus a clear asymmetry between the studies addressing effective health coverage and the burden of disease.

Utilisation and quality are the key parameters to measure effective coverage. Despite this, roughly a fifth of all assessments failed to reach an adequate measurement of utilisation conditioned to a normative need, while another fifth failed even to have any measure of quality. Among the assessments that included a quality parameter, few were based on the concept of health gains proposed by Shengelia et al. [43, 50-53]. This confirms what had been mentioned previously, that quality is the most difficult and complex component to be obtained for effective coverage measures.

Ng et al. classified the quality parameter of effective coverage in six categories, according to the approach used to calculate it [33]: 1) content of care, 2) biomarkers, 3) cohort

registration, 4) exposure matching, 5) statistical methods and 6) risk-adjusted outcomes. I believe this categorisation is neither practical nor useful to understanding the properties of the indicator of effective coverage. For example, we identified several studies where the health benefit of the intervention was detected biologically through biomarkers, according to Ng et al.'s categories, such as BP for hypertension or antibodies for vaccination. The size of the benefit in those cases can be estimated using instrumental variables through statistical methods, according to Ng et al., which can be applied to cross-sectional data or a cohort study. This shows that the categories proposed by Ng et al. are not mutually exclusive. In addition, some quality measures can be difficult to classify in any of the Ng et al. categories. For example, the quality criteria for interventions addressing vision problems [43] could be difficult to assign to any of the categories. This classification also does not inform about the intrinsic properties of the effective coverage estimator.

However, my proposal based on the ICF-2 plus three additional categories allow us to appreciate how close or far away the quality parameter is from the concept of health gain. In relation to this, the most useful parameter of quality is that measured in terms of health status or disability (as originally proposed by Shengelia et al.), because it allows comparisons and even adds up to the effective coverage from different interventions to summarise the performance of a whole health system. Only one of the 246 assessments included in this review used such an approach, showing how far the empirical use of effective coverage is from its theoretical optimum. All measures of quality based on contents of care (e.g., a pregnant woman receives at least 4 antenatal visits) or a quality index of a health facility (e.g., the availability of ultra-sonographer and at least one skilled professional) are merely proxies of potential health gain.

The concept of quality proposed by Shengelia et al.[32] can be easily misleading and is confounded by the concept of effectiveness, particularly when the quality parameter is used to calculate the 'effective' part of the coverage. That is evident in the scoping review by Jannait et al. [54], where the authors used the terms effectiveness and quality indistinctly. However, according to Shengelia et al., quality is the ratio between the actual health gain and the maximum health gain under optimum conditions [32]. The difference between the actual and optimum scenario would be given by the provider conditions and also by the users through adherence. For effectiveness, we compare the benefit observed among those receiving the intervention with those who are not receiving such intervention.

The five types of coverage proposed by Tanahashi were frequently cited by the selected studies, but very few were explored. The exception was those studied that followed a bottleneck approach to investigate how to improve the performance of the healthcare intervention under study. The scarce research on availability, accessibility, affordability and acceptability can be due to the need to conjugate different approaches, study designs and sample frameworks, which is difficult to encompass in a single study.

1.4.3.3 Limitations

My review has several limitations. One of them was the strategy of search, which may have been low in sensitivity since the initial search found only 471 records. That number is substantially smaller than the number found in other reviews, where the search generally found more than 5,000 abstracts. Nonetheless, my analysis includes all the studies found previously and adds many more new studies than those included in previous reviews. Moreover, the authors of one of those reviews [54], explicitly declared to have excluded all studies that were not strictly based on the Shengelia et al. conceptual framework.

Another limitation was that the selection of abstracts and extraction of information was carried out by only one researcher. This may have increased the risk of misidentification and misclassification of information. However, for that reason, all data extracted was checked twice using several classification tables, which allowed me to verify consistency across them. I did not measure the quality of the selected studies as there was no intention to summarise any magnitude of effect extracted from them. Not measuring the quality of studies is consistent with the traditional criteria used in scoping reviews [36].

1.4.3.4. Conclusions

The results of this scoping review show the gaps between the theoretical framework of effective coverage and its practical use in research. Even though effective coverage is an old metric that has recently been applied more consistently, it still seems that it should be improved to ensure its use in a broader range of interventions and settings. The parameter of quality of the healthcare interventions, understood as the potential health gain, seems to be the most challenging aspect of this metric. Any progress on how to measure effective

coverage may have the potential to optimise the indicators used for tracking progress toward the goals of UHC and how to assess the performance of health systems.

1.5. Research question

This thesis aims to develop a new procedure to measure effective coverage, overcoming limitations observed in its use in the scientific literature that has been implemented to track progress in UHC.

This challenge can be translated into the following research question:

> *How can effective coverage be estimated practically, including the concept of potential health gains in agreement with its most recent conceptual framework (Shengelia et al.) across different healthcare interventions and countries?*

1.6. Research aims and objectives

The main aim of this thesis is to develop a practical procedure to estimate the effective coverage of healthcare interventions addressed to common health conditions, using the concept of potential health gains in agreement with the most recent conceptual framework of effective coverage (form Shengelia et al.). The result could contribute to guiding decision-making processes in achieving WHO-UHC goals.

The specific objectives are:

1. To determine to what extent the metric of effective coverage could be used in scientific literature and the gap between how it is usually measured and its most recent conceptual framework (from Shengelia et al.).
2. To develop a new approach for measuring effective coverage, including a quality component base on the concept of potential health gains attributable to health interventions.
3. To use the novel approach to determine the effective coverage across different countries and health interventions.

1.7. Structure of the thesis

Following this introduction, in chapter two, I assess the psychometric properties of a questionnaire that explores disability and health status using the framework of the ICF-2. I also describe how to calculate a latent variable for those constructs. In chapter three, I present a new approach to measuring attributional fractions on continuous variables. That procedure is applied to the latent variable of disability developed in chapter two to measure the burden of disability attributable to different diseases. Chapter four shows the new proposal to measure effective coverage, which is applied to the treatment of a specific health condition. The procedure is implemented using the latent variable of disability and the attributional fractions method developed in the previous chapters. Chapters two, three and four address the second specific objective. Finally, in chapter five, the new procedure to estimate effective coverage is used to assess a set of interventions against diseases in different countries (specific objective three). Chapter six provides a discussion about this new approach, its scope related to how to track progress in UHC and future challenges.

References of chapter one

1. McCarthy M: **A brief history of the World Health Organization**. *Lancet* 2002, **360**(9340):1111-1112.
2. Pandey KR: **From health for all to universal health coverage: Alma Ata is still relevant**. *Globalization and Health* 2018, **14**(1):62.
3. Gorsky M, Sirrs C: **The Rise and Fall of "Universal Health Coverage" as a Goal of International Health Politics, 1925-1952**. *Am J Public Health* 2018, **108**(3):334-342.
4. Assembly WH: **World Health Assembly Resolution 58.33. Sustainable health financing, universal coverage and social health insurance**. In: *5833.* Edited by Organization WH. Geneva; 2005.
5. World Health Organization: **Achieving Universal Health Coverage: Developing the Health Financing System.** In: *Technical Briefs for Policy-Makers* World Health Organization; 2005.
6. World Health Organization: **The world health report 2008 : primary health care now more than ever.** Switzerland; 2008.
7. World Health Organization: **The World Health Report. Health Systems Financing. The path to universal coverage**. Switzerland: World Health Organization; 2010.
8. O'Connell T, Rasanathan K, Chopra M: **What does universal health coverage mean?** *Lancet* 2014, **383**(9913):277-279.
9. Nations U: **General Assembly: Global health and foreign policy**. In: *A/67/L36* Edited by Nations U, vol. A/67/L.36 New York: United Nations; 2012.
10. Nations U: **Resolution adopted by the General Assembly on 25 September 2015: 70/1. Transforming our world: the 2030 Agenda for**

Sustainable Development. In: *A/RES/70/1.* Edited by Nations U, vol. A/RES/70/1. New York: United Nations; 2015.

11. Nations U: **Resolution adopted by the General Assembly on 6 July 2017**. In: *A/RES/71/313.* Edited by Nations U, vol. A/RES/71/313. New York: United Nations; 2017.
12. Tichenor M, Sridhar D: **Universal health coverage, health systems strengthening, and the World Bank**. *BMJ* 2017, **358**:j3347.
13. Boerma T, Eozenou P, Evans D, Evans T, Kieny M-P, Wagstaff A: **Monitoring Progress towards Universal Health Coverage at Country and Global Levels**. *PLOS Medicine* 2014, **11**(9).
14. Boerma T, AbouZah C, Evans D, Evans T: **Monitoring Intervention Coverage in the Context of Universal Health Coverage**. *PLOS Medicine* 2014, **11**(9).
15. World Health Organization: **Tracking universal health coverage: first global monitoring report**. In. Switzerland; 2015.
16. World Health Organization and the International Bank for Reconstruction and Development / The World Bank: **Tracking universal health coverage: 2017 global monitoring report**. Switzerland.; 2017.
17. Shiffman J, Shawar YR: **Strengthening accountability of the global health metrics enterprise**. *Lancet (London, England)* 2020, **395**(10234):1452-1456.

18. World Health Organization: **The World Health Report 2000. Health Systems: Improving Performance.** : World Health Organization; 2000.
19. Bank W: **World Development Report 1993. Investing in Health**: Oxford University Press; 1993.
20. Fullman N, Barber RM, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, Abdulkader RS, Abdulle AM, Abera SF *et al*: **Measuring progress and projecting attainment on the basis of past trends of the health-related Sustainable Development Goals in 188 countries: an analysis from the Global Burden of Disease Study 2016**. *The Lancet* 2017, **390**(10100):1423-1459.
21. Nolte E, McKee M: **Measuring the health of nations: analysis of mortality amenable to health care**. *Bmj* 2003, **327**(7424):1129.
22. Charlton JR, Hartley RM, Silver R, Holland WW: **Geographical variation in mortality from conditions amenable to medical intervention in England and Wales**. *Lancet* 1983, **1**(8326 Pt 1):691-696.
23. Ng M, Fullman N, Dieleman JL, Flaxman AD, Murray CJL, Lim SS: **Effective Coverage: A Metric for Monitoring Universal Health Coverage**. *PLOS Medicine* 2015, **11**(9).
24. Szwarcwald CL: **On the World Health Organisation's measurement of health inequalities**. *Journal of Epidemiol Community Health* 2002, **56**.
25. Blendon RJ, Kim M, Benson JM: **The Public Versus The World Health Organization On Health System Performance**. *Health Affairs*, **20**(3):10-20.
26. Scientific Peer Review Group: **Report of the Scientific Peer Review Group on Health Systems Performance Assesment**. In.: World Health Organization; 2002.
27. World Health Organization: **Health Systems Performance Assessment. Debates, Methods and Empiricism**. Geneva: World Health Organization; 2003.
28. Murray CJL: **Assessing health systems performance using information on effective coverage of interventions**. In: *Theme: Building health information systems.* Australia: The University of Queensland; 2009: 22.
29. European Observatory on Health Systems and Policies: **Health system performance comparison. An agenda for policy, information and research**: Open University Press; 2013.
30. Tanahashi T: **Health service coverage and its evaluation**. *Bulletin of the World Health Organization* 1978, **56**(2):295-303.
31. Richard C, Roy S, Caroline G: **Jonathan Bradshaw on Social Policy. Selecting Writings 1972 - 2011**. York, United Kingdom: University of York; 2013.
32. Shengelia B, Tandon A, B. Adams O, Murray CJL: **Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy**. *Social Science & Medicine* 2005, **61**:97-109.
33. Ng M, Fullman N, Dieleman JL, Flaxman AD, Murray CJ, Lim SS: **Effective coverage: a metric for monitoring Universal Health Coverage**. *PLoS Med* 2014, **11**(9):e1001730.
34. De Silva MJ, Lee L, Fuhr DC, Rathod S, Chisholm D, Schellenberg J, Patel V: **Estimating the coverage of mental health programmes: a systematic review**. *International journal of epidemiology* 2014, **43**(2):341-353.

35. Mathers CD, Murray CJM, Lopez AD, Solomon JA, Sadana R, Tandon A: **Estimates of healthy life expectancy for 191 countries in the year 2000: methods and results**. In: *Global Programme on Evidence for Health Policy Discussion Paper* World Health Organization; 2001.
36. Arksey H, O'Malley L: **Scoping studies: towards a methodological framework**. *International Journal of Social Research Methodology* 2005, **8**(1):19-32.
37. The PME: **The PLOS "Monitoring Universal Health Coverage" Collection: Managing Expectations**. *PLOS Medicine* 2014, **11**(9):e1001732.
38. **PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation**. *Annals of Internal Medicine* 2018, **169**(7):467-473.
39. World Health Organization: **International Classification of Functioning, Disability and Health (ICF)**: World Health Organization; 2001.
40. Salomon JA, Mathers CD, Chatterji S, Sadana R, Üstün TB, Murray CJL: **Quantifying Individual Levels of Heath: Definitions, Concepts, and MEasurement Issues**. In: *Health Systems Performance Assessment Debates, Methods and Empiricism.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 301-318.
41. De Silva MJ, Lee L, Fuhr DC, Rathod S, Chisholm D, Schellenberg J, Patel V: **Estimating the coverage of mental health programmes: a systematic review**. *Int J Epidemiol* 2014, **43**(2):341-353.
42. Amouzou A, Leslie HH, Ram M, Fox M, Jiwani SS, Requejo J, Marchant T, Munos MK, Vaz LME, Weiss W *et al*: **Advances in the measurement of coverage for RMNCH and nutrition: from contact to effective coverage**. *BMJ global health* 2019, **4**:e001297.
43. Lozano R, Soliz P, Gakidou E, Abbott-Klafter J, Feehan DM, Vidal C, Ortiz JP, Murray CJL: **Benchmarking of performance of Mexican states with effective coverage**. *The Lancet* 2006, **368**(9548):1729-1741.
44. Andrews G, Sanderson K, Corry J, Lapsley HM: **Using epidemiological data to model efficiency in reducing the burden of depression***. *The journal of mental health policy and economics* 2000, **3**(4):175-186.
45. Health Do: **IAPT three-year report. The first million patients**. In. England: Department of Health; 2012.
46. Viviescas-Vargas DP, Idrovo AJ, Lopez-Lopez E, Uicab-Pool G, Herrera-Trujillo M, Balam-Gomez M, Hidalgo-Solorzano E: **[Effective coverage to manage domestic violence against women in Mexican municipalities: limits of metrics]**. *Cobertura efectiva del manejo de la violencia contra mujeres en municipios Mexicanos: limites de la metrica* 2013, **47**(4):781-787.
47. Liu YL, Rao KQ, Wu J, Gakidou E: **Health system reform in China 7: China's health system performance**. *Lancet (British edition)* 2008, **372**(9653):1914-1923.
48. Just MR, Carden SW, Li S, Baker KK, Gambhir M, Fung IC-H: **The impact of shared sanitation facilities on diarrheal diseases with and without an environmental reservoir: a modeling study**. *Pathogens and global health* 2018, **112**(4):195-202.
49. **Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries**

and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018, **392**(10159):1859-1922.

50. Guerrero-Nunez S, Valenzuela-Suazo S, Cid-Henriquez P: **Effective Universal Coverage of Diabetes Mellitus Type 2 in Chile**. *Revista latino-americana de enfermagem* 2017, **25**:e2871.
51. Jordans MJD, Luitel NP, Kohrt BA, Rathod SD, Garman EC, De Silva M, Komproe IH, Patel V, Lund C: **Community-, facility-, and individual-level outcomes of a district mental healthcare plan in a low-resource setting in Nepal: A population-based evaluation**. *PLoS medicine* 2019, **16**(2):e1002748.
52. Aguilera X, Castillo-Laborde C, Ferrari MN, Delgado I, Ibanez C: **Monitoring and evaluating progress towards universal health coverage in Chile**. *PLoS Med* 2014, **11**(9):e1001676.
53. Huda T, Khan JA, Ahsan KZ, Jamil K, El Arifeen S: **Monitoring and evaluating progress towards Universal Health Coverage in Bangladesh**. *PLoS Med* 2014, **11**(9):e1001722.
54. Jannati A, Sadeghi V, Imani A, Saadati M: **Effective coverage as a new approach to health system performance assessment: a scoping review**. *BMC health services research* 2018, **18**(1):886.
55. Cesar CL, Walker GJ: **Diversity in provision and utilization of maternal and child health care in an urban area of Brazil**. *Annals of tropical paediatrics* 1986, **6**(3):167-174.
56. al-Sekait MA: **The traditional midwife in Saudian villages**. *Journal of the Royal Society of Health* 1989, **109**(4):137-140.
57. Borus PK: **Missed opportunities and inappropriately given vaccines reduce immunisation coverage in facilities that serve slum areas of Nairobi**. *East African medical journal* 2004, **81**(3):124-129.
58. Gakidou E, Lozano R, González-Pier E, Abbott-Klafter J, Barofsky JT, Bryson-Cahn C, Feehan DM, Lee DK, Hernández-Llamas H, Murray CJL: **Assessing the effect of the 2001–06 Mexican health reform: an interim report card**. *The Lancet* 2006, **368**(9550):1920-1935.
59. Gakidou E, Nordhagen S, Obermeyer Z: **Coverage of cervical cancer screening in 57 countries: low average levels and large inequalities**. *PLoS Medicine* 2008, **5**(6):e132.
60. Kumar P, Prajapati P, Saxena D, Kavishwar AB, Kurian G: **An evaluation of coverage and compliance of mass drug administration 2006 for elimination of lymphatic filariasis in endemic areas of gujarat**. *Indian journal of community medicine : official publication of Indian Association of Preventive & Social Medicine* 2008, **33**(1):38-42.
61. Hussein M, Jira C, Girma B: **Assessment of effective coverage of voluntary counseling and testing services in major health facilities of jimma zone, South ethiopia**. *Ethiopian journal of health sciences* 2011, **21**:9-18.
62. Hussein M, Jira C, Girma B: **Assessment of Effective Coverage of HIV Prevention of Pregnant Mother to Child Transimission Services in Jimma Zone, South West Ethiopia**. *Ethiopian journal of health sciences* 2011, **21**:1-7.
63. Idzerda L, Adams O, Patrick J, Schrecker T, Tugwell P: **Access to primary healthcare services for the Roma population in Serbia: a secondary data**

analysis. *BMC international health and human rights* 2011, **11**(101088678):10.

64. Martinez S, Carrasquilla G, Guerrero R, Gomez-Dantes H, Castro V, Arreola-Ornelas H, Bedregal P, Caribe RdIsCEdIdSdOdSpALy, Vidal C MS, Guerrero R, Carrasquilla G, Solano G, Rosello M, Evans R, Peraza J, Rica C, Kestler E, Lozano R, Mendez O, Dorantes J, Knaul FM, Lopez E, Gomez H, Victoria Castro M, Carcamo C, Quiterio G, Pulido P: **[Effective coverage of health interventions in Latin America and the Caribbean: metrics for the assessment of health systems performance]**. *Cobertura efectiva de las intervenciones en salud de America Latina y el Caribe: metrica para evaluar los sistemas de salud* 2011, **53**(404371):s78-84.
65. Lopez-Lopez E, Gutierrez-Soria D, Idrovo AJ: **Evaluation of a diabetes care program using the effective coverage framework**. *International journal for quality in health care : journal of the International Society for Quality in Health Care* 2012, **24**(6):619-625.
66. Mehta S, Shah V, Verma A, Patel NB, Bansal RK: **Comparison of coverage and compliance of Mass Drug Administration 2012 in Surat, India**. *National Journal of Community Medicine* 2012, **3**(3):468-472.
67. Ranganath T, Reddy NR: **Elimination of lymphatic filariasis: mass drug administration in endemic areas of (bidar district) karnataka-2008**. *Indian journal of community medicine : official publication of Indian Association of Preventive & Social Medicine* 2012, **37**(4):219-222.
68. Knaul FM, Arreola-Ornelas H, Méndez-Carniado O, Bryson-Cahn C, Barofsky J, Maguire R, Miranda M, Sesma S: **Evidence is good for your health system: policy reform to remedy catastrophic and impoverishing health spending in Mexico**. *The Lancet* 2006, **368**(9549):1828-1841.
69. Beer N, Ali AS, Shakely D, Elfving K, Al-Mafazy A-WH, Msellem M, Petzold M, Bjorkman A, Kallander K: **High effective coverage of vector control interventions in children after achieving low malaria transmission in Zanzibar, Tanzania**. *Malaria journal* 2013, **12**(101139802):38.
70. Ghosh S, Samanta A, Kole S: **Mass drug administration for elimination of lymphatic filariasis: Recent experiences from a district of West Bengal, India**. *Tropical parasitology* 2013, **3**(1):67-71.
71. Gutierrez JP: **[Gaps in effective coverage by socioeconomic status and poverty condition]**. *Brechas en cobertura efectiva por nivel socioeconomico y condicion de pobreza* 2013, **55**(404371):S106-111.
72. Nesbitt RC, Lohela TJ, Manu A, Vesel L, Okyere E, Edmond K, Owusu-Agyei S, Kirkwood BR, Gabrysch S: **Quality along the continuum: a health facility assessment of intrapartum and postnatal care in Ghana**. *PloS one* 2013, **8**(11):e81089.
73. Roy RN, Sarkar AP, Misra R, Chakroborty A, Mondal TK, Bag K: **Coverage and awareness of and compliance with mass drug administration for elimination of lymphatic filariasis in Burdwan District, West Bengal, India**. *Journal of health, population, and nutrition* 2013, **31**(2):171-177.
74. Scheibe FJB, Waiswa P, Kadobera D, Muller O, Ekstrom AM, Sarker M, Neuhann HWF: **Effective coverage for antiretroviral therapy in a Ugandan district with a decentralized model of care**. *PloS one* 2013, **8**(7):e69433.

75. Shah H, Chaudhari VP, Desai B, Kantharia SL: **Data validation of immunization under RCH programme among poor performing blocks of Surat district**. *Healthline, Journal of Indian Association of Preventive and Social Medicine* 2013, **4**(1):70-73.
76. Singh S, Patel M, Kushwah SS: **An evaluation of mass drug administration compliance against filariasis of tikamgarh district of madhya pradesh-a household-based community study**. *Journal of family medicine and primary care* 2013, **2**(2):178-181.
77. Adhikari SR: **Towards universal health coverage: an example of malaria intervention in Nepal**. *WHO South-East Asia journal of public health* 2014, **3**(1):103-112.
78. Pirkis J, Ftanou M, Williamson M, Machlin A, Spittal MJ, Bassilios B, Harris M: **Australia's Better Access initiative: an evaluation**. *Aust N Z J Psychiatry* 2011, **45**(9):726-739.
79. Araya R, Zitko P, Markkula N: **The Impact of Universal Health Care Programmes on Improving 'Realized Access' to Care for Depression in Chile**. *Adm Policy Ment Health* 2018, **45**(5):790-799.
80. Aagaard J, ielsen JA: **Experience from the first ACT programme in Denmark. I. Baseline evaluation 1981-2000**. *Nord J Psychiatry* 2004, **58**(2):165-169.
81. Lin C, Wu Z, Rou K, Yin W, Wang C, Shoptaw S, Detels R: **Structural-level factors affecting implementation of the methadone maintenance therapy program in China**. *J Subst Abuse Treat* 2010, **38**(2):119-127.
82. Martini P, Cecchini M, Corlito G, D'Arco A, Nascimbeni P: **A model of a single comprehensive mental health service for a catchment area: a community alternative to hospitalization**. *Acta psychiatrica Scandinavica Supplementum* 1985, **316**:95-120.
83. Marinoni A, Torre E, Allegri G, Comelli M: **Lomest psychiatric case register: the statistical context required for planning**. *Acta Psychiatr Scand* 1983, **67**(2):109-117.
84. Hodgins S, D'Agostino A: **The quality-coverage gap in antenatal care: toward better measurement of effective coverage**. *Global health, science and practice* 2014, **2**(2):173-181.
85. Kulkarni P, Kumar R, Rajegowda RM, Channabasappa HG, Ashok NC: **MDA program against lymphatic filariasis: are we on the path to success? Experience from Uttara Kannada district, Karnataka**. *International Journal of Medicine and Public Health* 2014, **4**(3):243-246.
86. Srivastava PK, Bhattacharjee J, Dhariwal AC, Krishnamoorthy K, Dash AP: **Elimination of lymphatic filariasis - current status and way ahead**. *Journal of Communicable Diseases* 2014, **46**(2):85-94.
87. Zhou G, Li JS, Ototo EN, Atieli HE, Githeko AK, Yan G: **Evaluation of universal coverage of insecticide-treated nets in western Kenya: field surveys**. *Malaria journal* 2014, **13**(101139802):351.
88. Baker U, Okuga M, Waiswa P, Manzi F, Peterson S, Hanson C, group Es: **Bottlenecks in the implementation of essential screening tests in antenatal care: Syphilis, HIV, and anemia testing in rural Tanzania and Uganda**. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 2015, **130**:S43-50.

89. Baker U, Peterson S, Marchant T, Mbaruku G, Temu S, Manzi F, Hanson C: **Identifying implementation bottlenecks for maternal and newborn health interventions in rural districts of the United Republic of Tanzania**. *Bulletin of the World Health Organization* 2015, **93**(6):380-389.
90. Colson KE, Zuniga-Brenes P, Rios-Zertuche D, Conde-Glez CJ, Gagnier MC, Palmisano E, Ranganathan D, Usmanova G, Salvatierra B, Nazar A *et al*: **Comparative Estimates of Crude and Effective Coverage of Measles Immunization in Low-Resource Settings: Findings from Salud Mesoamerica 2015**. *PloS one* 2015, **10**(7):e0130697.
91. Crowe S, Utley M, Walker G, Panovska-Griffiths J, Grove P, Pagel C: **A novel approach to evaluating the UK childhood immunisation schedule: estimating the effective coverage vector across the entire vaccine programme**. *BMC infectious diseases* 2015, **15**(100968551):585.
92. Engle-Stone R, Nankap M, Ndjebayi AO, Vosti SA, Brown KH: **Estimating the Effective Coverage of Programs to Control Vitamin A Deficiency and Its Consequences Among Women and Young Children in Cameroon**. *Food and nutrition bulletin* 2015, **36**(3):S149-171.
93. Galactionova K, Tediosi F, de Savigny D, Smith T, Tanner M: **Effective coverage and systems effectiveness for malaria case management in sub-Saharan African countries**. *PloS one* 2015, **10**(5):e0127818.
94. Marathe N, Chalisgaonkar C: **Mass drug administration coverage evaluation for elimination of lymphatic filariasis in Chhatarpur district of Madhya Pradesh**. *International Journal of Medical Science and Public Health* 2015, **4**(7):927-932.
95. Mishra A, Trivedi R, Sharma D, Niranjan A, Sharma S: **Mid-term assessment of mass drug administration of DEC for filariasis in Rewa district of Madhya Pradesh**. *International Journal of Medical Science and Public Health* 2015, **4**(4):549-553.
96. Aaron GJ, Strutt N, Boateng NA, Guevarra E, Siling K, Norris A, Ghosh S, Nyamikeh M, Attiogbe A, Burns R *et al*: **Assessing Program Coverage of Two Approaches to Distributing a Complementary Feeding Supplement to Infants and Young Children in Ghana**. *PloS one* 2016, **11**(10):e0162462.
97. H. L, C.P. S, K.H. B, R. E-S: **Predicted effects of current and potential micronutrient intervention programs on adequacy of folate and vitamin B-12 intake in a national sample of women and young children in Cameroon**. *FASEB Journal* 2016, **30**.
98. J.-L. K, R. N, H. H, J. R, A. S: **Quality of care and effective coverage of child health services in rural Burkina Faso**. *Revue d'Epidemiologie et de Sante Publique* 2016, **64**:S242.
99. Leyvraz M, Rohner F, Konan AG, Esso LJCE, Woodruff BA, Norte A, Adiko AF, Bonfoh B, Aaron GJ: **High Awareness but Low Coverage of a Locally Produced Fortified Complementary Food in Abidjan, Cote d'Ivoire: Findings from a Cross-Sectional Survey**. *PloS one* 2016, **11**(11):e0166295.
100. Nguyen M, Poonawala A, Leyvraz M, Berger J, Schofield D, Nga TT, Van TK, Hoa DTB, Wieringa FT: **A Delivery Model for Home Fortification of Complementary Foods with Micronutrient Powders: Innovation in the Context of Vietnamese Health System Strengthening**. *Nutrients* 2016, **8**(5).

101. Patel V, Xiao S, Chen H, Hanna F, Jotheeswaran AT, Luo D, Parikh R, Sharma E, Usmani S, Yu Y *et al*: **The magnitude of and health system responses to the mental health treatment gap in adults in India and China**. *Lancet (London, England)* 2016, **388**(10063):3074-3084.
102. S. B, V. C, L. C: **Barriers in access to health services for people living with HIV in Moldova**. *Journal of the International AIDS Society* 2016, **19**:87.
103. Servan-Mori E, Avila-Burgos L, Nigenda G, Lozano R: **A Performance Analysis of Public Expenditure on Maternal Health in Mexico**. *PloS one* 2016, **11**(4):e0152635.
104. Yawson AE, Awoonor-Williams JK, Sagoe-Moses I, Aboagye PK, Yawson AO, Senaya LK, Bonsu G, Eleeza JB, Agongo EEA, Banskota HK: **Bottleneck analysis approach to accelerate newborn care services in two regions in Ghana: implications for national newborn care**. *Public health* 2016, **141**:245-254.
105. Comfort AB, Krezanoski PJ: **The effect of price on demand for and use of bednets: evidence from a randomized experiment in Madagascar**. *Health policy and planning* 2017, **32**(2):178-193.
106. Engle-Stone R, Perkins A, Clermont A, Walker N, Haskell MJ, Vosti SA, Brown KH: **Estimating Lives Saved by Achieving Dietary Micronutrient Adequacy, with a Focus on Vitamin A Intervention Programs in Cameroon**. *The Journal of nutrition* 2017, **147**(11):2194S-2203S.
107. H. L, C. S, S. V, J.G. A, K. B, R. E-S: **Simulations of the impacts of biofortified crops on vitamin A intake of young children and iron intake of women of reproductive age in Cameroon**. *Annals of Nutrition and Metabolism* 2017, **71**:777-778.
108. Khan NAM, Gattani PL, Inamdar IA, Domple VK, Nina MB: **Coverage evaluation and compliance of mass drug administration campaign in Nanded district of Maharashtra**. *National Journal of Community Medicine* 2017, **8**(11):684-687.
109. Kiwanuka Henriksson D, Fredriksson M, Waiswa P, Selling K, Swartling Peterson S: **Bottleneck analysis at district level to illustrate gaps within the district health system in Uganda**. *Global health action* 2017, **10**(1):1327256.
110. Larson E, Vail D, Mbaruku GM, Mbatia R, Kruk ME: **Beyond utilization: measuring effective coverage of obstetric care along the quality cascade**. *International journal for quality in health care : journal of the International Society for Quality in Health Care* 2017, **29**(1):104-110.
111. Leslie HH, Malata A, Ndiaye Y, Kruk ME: **Effective coverage of primary care services in eight high-mortality countries**. *BMJ global health* 2017, **2**(3):e000424.
112. Leyvraz M, Aaron GJ, Poonawala A, van Liere MJ, Schofield D, Myatt M, Neufeld LM: **Coverage of Nutrition Interventions Intended for Infants and Young Children Varies Greatly across Programs: Results from Coverage Surveys in 5 Countries**. *The Journal of nutrition* 2017, **147**(5):995S-1003S.
113. Lindtjorn B, Mitiku D, Zidda Z, Yaya Y: **Reducing Maternal Deaths in Ethiopia: Results of an Intervention Programme in Southwest Ethiopia**. *PloS one* 2017, **12**(1):e0169304.
114. Nguhiu PK, Barasa EW, Chuma J: **Determining the effective coverage of maternal and child health services in Kenya, using demographic and health**

survey data sets: tracking progress towards universal health coverage. *Tropical medicine & international health : TM & IH* 2017, **22**(4):442-453.

115. R. G, M. N, R. S, M. B, R.V. D, C.R. M, S. M, B. P, M. R, B. T *et al*: **Identifying gaps in the continuum of care for hypertension and diabetes in two Indian communities**. *BMC health services research* 2017, **17**(1):846.
116. Rios-Blancas MJ, Cahuana-Hurtado L, Lamadrid-Figueroa H, Lozano R: **[Effective coverage of treatment of hypertension in Mexican adults by states]**. *Cobertura efectiva del tratamiento de la hipertension arterial en adultos en Mexico por entidad federativa* 2017, **59**(2):154-164.
117. S.A. V, R. E-S, H. L, J. K, A. C, N. W: **From dietary adequacy for vitamin A in children and folate in women to children's lives saved in Cameroon: A bioeconomic model to inform cost-effective policy choices**. *FASEB Journal* 2017, **31**(1).
118. Sharma J, Leslie HH, Kundu F, Kruk ME: **Poor Quality for Poor Women? Inequities in the Quality of Antenatal and Delivery Care in Kenya**. *PloS one* 2017, **12**(1):e0171236.
119. Arredondo A, Azar A, Recaman AL: **Challenges and dilemmas on universal coverage for non-communicable diseases in middle-income countries: evidence and lessons from Mexico**. *Globalization and health* 2018, **14**(1):89.
120. Arsenault C, Jordan K, Lee D, Dinsa G, Manzi F, Marchant T, Kruk ME: **Equity in antenatal care quality: an analysis of 91 national household surveys**. *The Lancet Global health* 2018, **6**(11):e1186-e1195.
121. Banerjee S, Ray S, Bhattacharya T, Naskar S, Mandal S, Das DK: **Mass drug administration coverage evaluation survey for lymphatic filariasis: an experience from Paschim Bardhaman district, West Bengal**. *Journal of Communicable Diseases* 2018, **50**(2):25-29.
122. Bhatia V, Giri PP, Sahoo SS, Preeti PS, Sahu DP: **Mass drug administration (MDA) for elimination of lymphatic filariasis: experiences from Nayagarh district of Odisha, India**. *Indian Journal of Community Health* 2018, **30**(3):287-292.
123. Brenner S, Mazalale J, Wilhelm D, Nesbitt RC, Lohela TJ, Chinkhumba J, Lohmann J, Muula AS, De Allegri M: **Impact of results-based financing on effective obstetric care coverage: evidence from a quasi-experimental study in Malawi**. *BMC health services research* 2018, **18**(1):791.
124. Carter ED, Ndhlovu M, Eisele TP, Nkhama E, Katz J, Munos M: **Evaluation of methods for linking household and health care provider data to estimate effective coverage of management of child illness: results of a pilot study in Southern Province, Zambia**. *Journal of global health* 2018, **8**(1):010607.
125. Charoendee K, Sriratanaban J, Aekplakorn W, Hanvoravongchai P: **Assessment of population coverage of hypertension screening in Thailand based on the effective coverage framework**. *BMC health services research* 2018, **18**(1):208.
126. Jannati A, Sadeghi V, Imani A, Saadati M: **Effective coverage as a new approach to health system performance assessment: a scoping review**. *BMC Health Serv Res* 2018, **18**(1):886.
127. Randive BB, Chaturvedi SD, Diwan V, Costa AD: **Effective coverage of institutional deliveries under the Janani Suraksha Yojana programme in high**

maternal mortality provinces of India: analysis of data from an annual health survey. *The Lancet* 2013, **381**:S123.

128. Travassos MA, Beyene B, Adam Z, Campbell JD, Mulholland N, Diarra SS, Kassa T, Oot L, Sequeira J, Reymann M *et al*: **Immunization Coverage Surveys and Linked Biomarker Serosurveys in Three Regions in Ethiopia**. *PLOS ONE* 2016, **11**(3):e0149970.
129. Colson KE, Gagnier MC, Palmisano E, Ranganathan D, Zúñiga-Brenes P, Ríos-Zertuche D, Ibarraran P, Murphy T, Anderson BW, Salvatierra B *et al*: **Comparative estimates of immunisation coverage from three different sources: results from the SM2015 evaluation**. *The Lancet* 2013, **381**:S32.
130. Koulidiati J-L, Nesbitt RC, Ouedraogo N, Hien H, Robyn PJ, Compaore P, Souares A, Brenner S: **Measuring effective coverage of curative child health services in rural Burkina Faso: a cross-sectional study**. *BMJ open* 2018, **8**(5):e020423.
131. Leyvraz M, David-Kigaru DM, Macharia-Mutie C, Aaron GJ, Roefs M, Tumilowicz A: **Coverage and Consumption of Micronutrient Powders, Fortified Staples, and Iodized Salt Among Children Aged 6 to 23 Months in Selected Neighborhoods of Nairobi County, Kenya**. *Food and nutrition bulletin* 2018, **39**(1):107-115.
132. Mahdavi M, Vissers J, Elkhuizen S, van Dijk M, Vanhala A, Karampli E, Faubel R, Forte P, Coroian E, van de Klundert J: **The relationship between context, structure, and processes with outcomes of 6 regional diabetes networks in Europe**. *PloS one* 2018, **13**(2):e0192599.
133. Murphy GAV, Gathara D, Mwachiro J, Abuya N, Aluvaala J, English M, Group HStDfNE, Ochola S AR, Wasunna A, Were F, Mutinda C, Maina B, Mutiso C, Githanga D, Kimutai D, Musoke R, Ochieng R, Macharia W, Nyamai R: **Effective coverage of essential inpatient care for small and sick newborns in a high mortality urban setting: a cross-sectional study in Nairobi City County, Kenya**. *BMC medicine* 2018, **16**(1):72.
134. Stones W, Peno WW: **Access to antenatal blood pressure measurement in Malawi: Findings from a national census of health facilities**. *Malawi medical journal : the journal of Medical Association of Malawi* 2018, **30**(3):141-145.
135. Willey B, Waiswa P, Kajjo D, Munos M, Akuze J, Allen E, Marchant T: **Linking data sources for measurement of effective coverage in maternal and newborn health: what do we learn from individual- vs ecological-linking methods?** *Journal of global health* 2018, **8**(1):010601.
136. Thapa K, Sanghvi H, Rawlins B, Karki YB, Regmi K, Aryal S, Aryal Y, Murakami P, Bhattarai J, Suhowatsky S: **Coverage, compliance, acceptability and feasibility of a program to prevent pre-eclampsia and eclampsia through calcium supplementation for pregnant women: an operations research study in one district of Nepal**. *BMC Pregnancy Childbirth* 2016, **16**(1):241.
137. Deming MS, Roungou JB, Kristiansen M, Heron I, Yango A, Guenengafo A, Ndamobissi R: **Tetanus toxoid coverage as an indicator of serological protection against neonatal tetanus**. *Bull World Health Organ* 2002, **80**(9):696-703.

138. Aaron GJ, Sodani PR, Sankar R, Fairhurst J, Siling K, Guevarra E, Norris A, Myatt M: **Household Coverage of Fortified Staple Food Commodities in Rajasthan, India**. *PLoS One* 2016, **11**(10):e0163176.
139. Aaron GJ, Friesen VM, Jungjohann S, Garrett GS, Neufeld LM, Myatt M: **Coverage of Large-Scale Food Fortification of Edible Oil, Wheat Flour, and Maize Flour Varies Greatly by Vehicle and Country but Is Consistently Lower among the Most Vulnerable: Results from Coverage Surveys in 8 Countries**. *J Nutr* 2017, **147**(5):984s-994s.
140. Agha S, Tappis H: **The timing of antenatal care initiation and the content of care in Sindh, Pakistan**. *BMC Pregnancy Childbirth* 2016, **16**(1):190.
141. Gebremedhin S, Samuel A, Mamo G, Moges T, Assefa T: **Coverage, compliance and factors associated with utilization of iron supplementation during pregnancy in eight rural districts of Ethiopia: a cross-sectional study**. *BMC Public Health* 2014, **14**:607.
142. Hayford KT, Shomik MS, Al-Emran HM, Moss WJ, Bishai D, Levine OS: **Measles vaccination coverage estimates from surveys, clinic records, and immune markers in oral fluid and blood: a population-based cross-sectional study**. *BMC Public Health* 2013, **13**:1211.
143. Heredia-Pi I, Servan-Mori E, Darney BG, Reyes-Morales H, Lozano R: **Measuring the adequacy of antenatal health care: a national cross-sectional study in Mexico**. *Bull World Health Organ* 2016, **94**(6):452-461.
144. Kanyangarara M, Munos MK, Walker N: **Quality of antenatal care service provision in health facilities across sub-Saharan Africa: Evidence from nationally representative health facility assessments**. *J Glob Health* 2017, **7**(2):021101.
145. Khan Z, Malik A, Pal SD, Khan MY: **Measles antibody titre in children up to 5 years of age in rural areas of Aligarh District, India**. *Saudi Med J* 2000, **21**(2):175-179.
146. Kyei NN, Chansa C, Gabrysch S: **Quality of antenatal care in Zambia: a national assessment**. *BMC Pregnancy Childbirth* 2012, **12**:151.
147. Leyvraz M, Wirth JP, Woodruff BA, Sankar R, Sodani PR, Sharma ND, Aaron GJ: **High Coverage and Utilization of Fortified Take-Home Rations among Children 6-35 Months of Age Provided through the Integrated Child Development Services Program: Findings from a Cross-Sectional Survey in Telangana, India**. *PLoS One* 2016, **11**(10):e0160814.
148. Marchant T, Tilley-Gyado RD, Tessema T, Singh K, Gautham M, Umar N, Berhanu D, Cousens S, Armstrong Schellenberg JR: **Adding content to contacts: measurement of high quality contacts for maternal and newborn health in Ethiopia, north east Nigeria, and Uttar Pradesh, India**. *PLoS One* 2015, **10**(5):e0126840.
149. Millar KR, McCutcheon J, Coakley EH, Brieger W, Ibrahim MA, Mohammed Z, Bassi A, Sambisa W: **Patterns and predictors of malaria care-seeking, diagnostic testing, and artemisinin-based combination therapy for children under five with fever in Northern Nigeria: a cross-sectional study**. *Malar J* 2014, **13**:447.
150. Mokdad AH, Gagnier MC, Colson KE, Dansereau E, Zúñiga-Brenes P, Ríos-Zertuche D, Haakenstad A, Johanns CK, Palmisano EB, Hernandez B *et al*:

Missed Opportunities for Measles, Mumps, and Rubella (MMR) Immunization in Mesoamerica: Potential Impact on Coverage and Days at Risk. *PLoS One* 2015, **10**(10):e0139680.

151. Nanthavong N, Black AP, Nouanthong P, Souvannaso C, Vilivong K, Muller CP, Goossens S, Quet F, Buisson Y: **Diphtheria in Lao PDR: Insufficient Coverage or Ineffective Vaccine?** *PLoS One* 2015, **10**(4):e0121749.
152. Ndyomugyenyi R, Katamanywa J: **Intermittent preventive treatment of malaria in pregnancy (IPTp): do frequent antenatal care visits ensure access and compliance to IPTp in Ugandan rural communities?** *Trans R Soc Trop Med Hyg* 2010, **104**(8):536-540.
153. Smith LA, Bruce J, Gueye L, Helou A, Diallo R, Gueye B, Jones C, Webster J: **From fever to anti-malarial: the treatment-seeking process in rural Senegal**. *Malar J* 2010, **9**:333.
154. Wangdi K, Gatton ML, Kelly GC, Clements AC: **Prevalence of asymptomatic malaria and bed net ownership and use in Bhutan, 2013: a country earmarked for malaria elimination**. *Malar J* 2014, **13**:352.
155. Bekuma TT, Mosisa Kebebew G, Desalegn Waktole Z, Markos Cafo J, Wirtu D, Gaddisa S: **Coverage assessment survey following trachoma mass drug administration (MDA) in six districts of Oromia, Western Ethiopia, 2017**. *PLoS neglected tropical diseases* 2019, **13**(12):e0007924.
156. Eboreime EA, Nxumalo N, Ramaswamy R, Ibisomi L, Ihebuzor N, Eyles J: **Effectiveness of the Diagnose-Intervene- Verify-Adjust (DIVA) model for integrated primary healthcare planning and performance improvement: an embedded mixed methods evaluation in Kaduna state, Nigeria**. *BMJ open* 2019, **9**(3):e026016.
157. Hashiguchi K, Nagata T, Mori K, Nagata M, Fujino Y, Ito M: **Occupational Health Services Improve Effective Coverage for Hypertension and Diabetes Mellitus at Japanese Companies**. *Journal of UOEH* 2019, **41**(3):271-282.
158. Leslie HH, Doubova SV, Perez-Cuevas R: **Assessing health system performance: effective coverage at the Mexican Institute of Social Security**. *Health policy and planning* 2019, **34**:ii67-ii76.
159. M.A. C, D. I, K.E. H, A. J, P.-E. D, J.-B. M, A.M. D, B. F, E. P, K. B *et al*: **Results of a pilot of targeted mass drug administration with sulfadoxine-pyrimethamine and primaquine as a component of a malaria elimination package in Haiti**. *American Journal of Tropical Medicine and Hygiene* 2019, **101**(5):417.
160. Mulebeke R, Wanzira H, Bukenya F, Eganyu T, Collborn K, Elliot R, Van Geertruyden J-P, Echodu D, Yeka A: **Implementing population-based mass drug administration for malaria: experience from a high transmission setting in North Eastern Uganda**. *Malaria journal* 2019, **18**(1):271.
161. Ntambi J, Abdirahman MA, Nabiwemba D, Ghimire P, Majeed SE, Mohamed K, Desie S, Musvaire R, Volege M, Shoham J *et al*: **Bottleneck analysis for the integrated management of acute malnutrition services in Somalia**. *Special issue on the continuum of care for acute malnutrition* 2019(60):56-60.
162. P. K, J.J. T, J. D, M.R. NM, K. R: **Mass drug administration programme against lymphatic filariasis-an evaluation of coverage and compliance in a northern Karnataka district, India**. *Clinical Epidemiology and Global Health* 2019.

163. Servan-Mori E, Cerecero-Garcia D, Heredia-Pi IB, Pineda-Antunez C, Sosa-Rubi SG, Nigenda G: **Improving the effective maternal-child health care coverage through synergies between supply and demand-side interventions: evidence from Mexico**. *Journal of global health* 2019, **9**(2):020433.
164. Smith TA, Pemberton-Ross P, Penny MA, Chitnis N: **Resurgence of malaria infection after mass treatment: a simulation study**. *Malaria journal* 2019, **18**(1):409.
165. Tsiachristas A, Gathara D, Aluvaala J, Chege T, Barasa E, English M: **Effective coverage and budget implications of skill-mix change to improve neonatal nursing care: an explorative simulation study in Kenya**. *BMJ global health* 2019, **4**(6):e001817.
166. Venkateswaran M, Bogale B, Abu Khader K, Awwad T, Friberg IK, Ghanem B, Hijaz T, Morkrid K, Froen JF: **Effective coverage of essential antenatal care interventions: A cross-sectional study of public primary healthcare clinics in the West Bank**. *PloS one* 2019, **14**(2):e0212635.
167. Wang W, Mallick L, Allen C, Pullum T: **Effective coverage of facility delivery in Bangladesh, Haiti, Malawi, Nepal, Senegal, and Tanzania**. *PloS one* 2019, **14**(6):e0217853.
168. Yakob B, Gage A, Nigatu TG, Hurlburt S, Hagos S, Dinsa G, Bowser D, Berman P, Kruk ME, Tekle E: **Low effective coverage of family planning and antenatal care services in Ethiopia**. *International journal for quality in health care : journal of the International Society for Quality in Health Care* 2019, **31**(10):725-732.
169. Fuseini G, Ismail HM, von Fricken ME, Weppelmann TA, Smith J, Ellis Logan RA, Oladepo F, Walker KJ, Phiri WP, Paine MJI *et al*: **Improving the performance of spray operators through monitoring and evaluation of insecticide concentrations of pirimiphos-methyl during indoor residual spraying for malaria control on Bioko Island**. *Malaria journal* 2020, **19**(1):35.
170. M. M, M. P, B. M, M. M, N. A, M. Y, P. M, S. D, N. R, R. H *et al*: **Insight into blood pressure targets for universal coverage of hypertension services in Iran: the 2017 ACC/AHA versus JNC 8 hypertension guidelines**. *BMC public health* 2020, **20**(1):347.
171. Sarma H, Mbuya MN, Tariqujjaman M, Rahman M, Askari S, Khondker R, Sultana S, Shahin SA, Bossert TJ, Banwell C *et al*: **Role of home visits by volunteer community health workers: to improve the coverage of micronutrient powders in rural Bangladesh**. *Public health nutrition* 2020(9808463):1-11.
172. Zhao Y, Mahal AS, Tang S, Haregu TN, Oldenburg B: **Effective coverage for hypertension treatment among middle-aged adults and the older population in China, 2011 to 2013: A nationwide longitudinal study**. *Journal of global health* 2020, **10**(1):010805.

Chapter 2: Psychometric properties of the Health State Description questionnaire: A proposal for a latent variable approach to health valuation process

2.1. Introduction

In the previous chapter, I stated that in order to estimate effective coverage following Shengelia et al.'s framework [1] it is necessary to quantify the quality of healthcare interventions. According to this conceptual framework, quality is understood as the relationship between the current health gain attributable to the healthcare, and the potential health gain assuming that people are receiving the optimum care. However, in the scoping review presented in the same chapter, I observed that almost none of the 246 effective coverage assessments under review used a direct measure of health gain. Instead, the assessments mostly used contents of care, a quality index from the health provider, or a biomarker, as a proxy of health gain. This is a limitation to applying Shengelia et al.'s framework and it is in agreement with statements from several different organisations, which point out the difficulties associated with quantifying the quality of healthcare [2, 3].

One of the most important challenges that this thesis attempts to overcome is estimating effective coverage using a direct measure of health gain. However, before developing the procedure, it is necessary to briefly reflect on what we understand by 'health' and to choose the most suitable theoretical framework regarding the methodological proposal of effective coverage. This reflection and the selected conceptual framework are developed in the second section of the chapter. In addition, different frameworks and procedures used to value health states are presented.

Since the approach that it is developed to measure health gains requires to be tested, in the third section, I will introduce the database I have chosen for that purpose. This database, taken from a survey, includes all the elements necessary for the selected theoretical

framework: a questionnaire to measure health status, variables required to determine the need of healthcare interventions and its coverage, and other covariates. In the same section, I will describe the methodology used to assess the psychometric properties of the health status questionnaire.

In the fourth section of the chapter, I will show the results of the analysis, while in the fifth part, I will discuss them. It will be argued that a latent variable of the health status can be used for valuating health states associated with diseases.

2.2. What we understand by 'health'

Measuring the health of populations is necessary to evaluate the effectiveness of health interventions, to assess the performance of health systems and monitoring purposes [4]. There are several summary measures of population health, usually combining fatal and non-fatal outcomes. These summary measures can be broadly classified as either those that measure health expectancies, such as active life expectancy (ALE), disability-adjusted life expectancy (DALE), or years of healthy life (YHL), among others; or, those that measure health gaps, such as the disability-adjusted life years (DALYs) or the healthy life years (HeaLYs) [5]. All these measures, within the non-fatal component of the metric, consider the time lived in different health states and the valuation of those states. The valuation of health states is known as 'weights' [6]. What we understand by health states and how such weights can be obtained is described below.

2.2.1 The concept of health

The definition of 'health state' is intimately related to the concept of health itself. In 1946, the WHO defined health as 'a state of complete physical, mental, and social wellbeing and not merely the absence of disease or infirmity' [7]. This broad definition, useful for political purposes, permeated importantly into the research on health metrics, especially during the first decades after the WHO constitution. Although several indices to measure health numerically were developed during this period [8], their analyses show the difficulties of operationalising the WHO's definition of health.

One of the common features in these first indices is the multidimensional concept of health, with each one exploring several domains of the functioning. For example, the Index of Independence of Activities of Daily Living created by Katz et al. (1963), inquires into bathing, dressing, toileting, transfer, continence and feeding [9]. Or, another example, the Cornell Medical index published in 1966 assesses 'stated symptoms or disorder' in several body systems (eyes and ears, respiratory, cardiovascular, musculoskeletal, among others), plus fatigability, frequency of illness, miscellaneous diseases, habits, inadequacy, depression, anxiety, sensitivity, anger and tension [10]. In addition the Health Index from Grogono and Woodgate (1971) explores the ability to work, hobbies and interactions, presence of malaise,

pain or suffering, communication, acts of daily living, eating and enjoying food, micturition and defecation, and sex life [11]. Meanwhile, the Breslow Index is based on disability, impairments, chronic conditions, symptoms, and energy levels [12]. These instruments highlight amongst many other, the plurality of health dimensions, as well as the lack of agreement about: which ones need to be explored (e.g., physical, mental), what type of dimensions (e.g., daily activities, symptoms, habits, diseases, impairments, etc.), and how they are acquired (e.g., including personal expectations or not) [13].

Criticism about the lack of a coherent concept of health can be seen in this first wave of instruments used to measure health. These critics also address the use of wellbeing as a concept equivalent to health, something that is observed in a number of the aforementioned indices. Several authors consider that health is only one among many contributors to wellbeing, such as education, environmental conditions, economic security, and so on[13].

As a response to critics about the lack of clarity in the concept of health, at least two important theoretical currents emerged during the seventies. Both were intended to define health, and both exerted a strong influence in later research. One of them was the descriptivist current, represented mainly by the work of Christopher Boorse, who defended the notion of health as the absence of disease. Under this approach, the disease is understood as a deviation from the normal functioning proper of the specie, and health, consequently, is defined as a functioning between certain ranges, usually delimited through some statistical approach (e.g., a tail of a distribution) [14, 15].

The other relevant theoretical current was the normativism, represented mainly by the work of Lennart Nordenfelt, who pointed out the distinction between illness and health, defining the latter from a positive perspective. Nordenfelt focused first on disability, but in order to provide a positive concept of health, he moved on to the use the 'semantic content of its positive contrary'; in other words, the 'ability'. For Nordenfelt, a healthy individual has 'the ability, given standard circumstances, to realise his vital goals, i.e. the set of goals which are necessary and together sufficient for his minimal happiness' [15, 16]. Nordenfelt also recognised a second-order of abilities, namely capabilities, which enable the ability to achieve a vital goal. Non-health is expressed as 'the lack of capabilities to produce the ability to achieve the vital goals' [15]. Proximities between the idea of the ability to achieve vital goals and the concept of human capabilities from Amartya Sen have been suggested [15]. However, the main critics of Nordenfelt's theoretical framework point to the lack of

description of the dimensions covered by vital goals, as well as the conditioning of the ability to realise vital goals to 'standard circumstances,' which would make any measure of health dependent on external criteria [15, 16].

Encompassing this theoretical progress in the conception of health, disease, and disability, in 1980 the WHO launched the first version of an International Classification of Impairments, Disability, and Handicaps [17]. This classification attempted to clarify different experiences related to health and sickness. In accordance with this conceptual model, an etiologic agent can cause changes in the structure or functioning of the body, which can or cannot make themselves evident by causing manifestations, such as symptoms and signs. The term impairment is referred to as a loss of normality in the structure or functions of the body. As a product of impairment, individuals can change their behaviour or ability to perform an activity (e.g. walking, dressing, listening, seeing, etc.), which is understood as a disability. Finally, the impairment or the disability, jointly or separately, can be a disadvantage to the individual relative to others in terms of being able to play different roles in their societies, which is described as a handicap (e.g. mobility, social integration, orientation, physical independence) [17]. However, some authors of this classification have recognised several shortcomings in the conceptualisations of this framework, especially related to the concept of a handicap [18, 19]. It is worth noting the proximity between the central idea of the 'ability to perform an activity' in the WHO classification, and the Nordenfelt concept of 'the ability to realise vital goals.'

From this period, between the '80s and '90s, come the best-known questionnaires used nowadays to measure health and related concepts, such as the EuroQol initiative (i.e. EQ-5D) from 1990, which attempts to measure health-related quality of life, to describe health states and value them [20]. Other similar approaches include the Short Form 36 Health Survey (SF-36) questionnaire, from 1992, which aims to measure health status [21], and the first version of the WHO - Disability Assessment Schedule (WHODAS-I) from 1999 [22, 23], among others [24, 25].

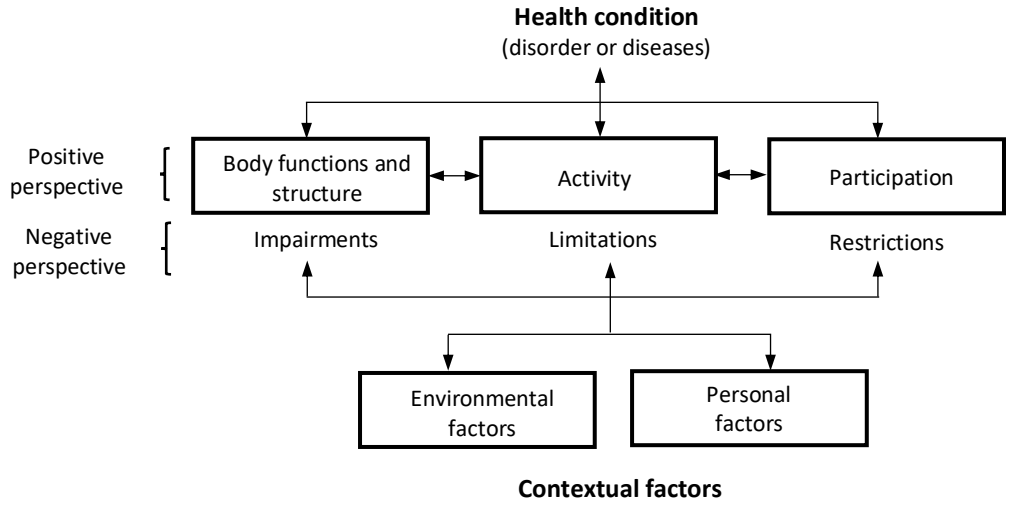
As a whole, the consensus reached during this period points to an idea that [13]: health is comprised of multidimensional states or conditions of functioning of the human body and mind; health is a different concept to well-being; and, health is an attribute of the individuals which can be aggregated to produce population level measurements.

2.2.2 An integrated framework for the concept of health: Functioning and disability

In line with this consensus on health, in 2001, the WHO presented the second International Classification of Functioning, Disability, and Health (ICF-2) [26], which replaced the former classification launched in 1980. This second version attempted to overcome some of the limitations of the previous version, providing a more solid conceptual framework for the constructs of health and disability. The conceptual framework of the ICF-2 is based on the idea that the functioning of individuals is underpinned by the *function* and *structure* of the body and mind, which can be expressed as the *capacity* to act or the *performance* of an action in different *domains* (see Figure 2.1). Actions can be expressed as *activities* (i.e. the execution of a task or action), or as *participation* (i.e. involvement in a life situation). Capacity is here understood to mean actions undertaken in a uniform or standard environment, while the performance refers to actions carried out in the own context of the individual. By definition, participation is always measured as performance because it involves life situations.

The domains of the functions and structures are classified into eight systems (e.g. nervous system/ mental functions, structures related to movement/ neuromusculoskeletal and movement-related functions, and so forth). Meanwhile, the domains of activities and participation include nine categories (learning and applying knowledge, general task and demands, communication, mobility, self-care, domestic life, interpersonal interaction and relationships, amongst others). While capacity is used to define a health state, it is understood as the capacity to exert actions in different areas or domains of functioning. At the same time, performance is used to describe disability, understood as the *limitations* and *restrictions* in performing actions in different areas of functioning [27]. In these terms, health and disability are not reciprocal concepts, and their difference lies in whether contextual factors are considered or not. As a consequence, the differences obtained in measuring capacity and performance are attributable to elements from the context. Contextual factors, according to the ICF-2 definitions, include those from an environmental or personal origin. When they act to promote participations, they are named *facilitators*, otherwise they are known as *barriers*.

Figure 2.1. Representation of the conceptual model of the International Classification of Functioning, Disability and Health (ICF-2), from the World Health Organization.



Adapted from [28]

Despite the conceptual distinction between capacity and performance, some authors have noticed the practical difficulties of building items for questionnaires in the context of population research, which allow differentiation between both concepts in the assessment of activities and participation [13]. Consequently, in several examples, no distinction is made between them, and the measurements of disability are simply treated as the complement of the health measurements (i.e. disability = 1- health).

Overall, the ICF-2's conceptual framework has several desirable attributes. First, it proposes a positive definition of health; that is, health is not understood as the absence of infirmity. Second, it recognises the multidimensional nature of health through the inclusion of domains of functioning. Third, the concept of health (the capacity to exert activities and participation) is aligned with previous normative definitions (the ability to perform an activity, and the ability to realize vital goals). Fourth, it offers a clear distinction between the measurements of functioning that do and do not consider contextual factors. This criterium allows a clear conceptual distinction between the concepts of health and disability, despite its practical difficulties in measuring them. Fifth, it provides a clear delimitation between the concepts of health/disability, and other concepts such as disease, wellbeing, and contextual factors.

Moreover, the creation of the ICF-2's conceptual framework coincided with a period of very intensive work at WHO, when several theoretical and metrical advances were developed and proposed. For example, the launching of this classification happened almost simultaneously with the presentation of the first study that evaluated the performance of health systems across the world [29], the first WHO report on the Global Burden of Disease including a large comparative risk assessment [30], the Multi-Country Survey and the World Health Surveys initiatives [31], and the first version of the second generation of the WHODAS questionnaire for measuring health and disability (WHODAS-II) [22].

The ICF-2's conceptual framework is still widely used in research, in medical practice and in statistical health records [32, 33]. Additionally, important public health initiatives - such as the current burden of disease studies from the Institute of Health Metrics and Evaluation (IHME) - adhere to this framework. In view of all these reasons using the CIF-2 definitions for health and disability in this thesis seems appropriate.

2.2.3 Health states

Since I have already clarified what is understood by health, in the context of this thesis, the next step is to point out what is understood by a health state. Health states are levels of functioning according to the body structures and functions, and the capacity or performance to exert activities and participation in different domains. Traditionally, people's health states are presented as a description such as: 'has lost one hand and part of the arm, leaving pain and tingling in the stump. The person needs help from others to lift objects or do daily activities such as cooking'; or 'feels persistent sadness and has lost interest in usual activities. The person sometimes sleeps badly, feels tired, or has trouble concentrating but still manages to function in daily life with extra effort': or 'I have some problems in walking about, I am unable to wash or dress myself, I have some problems with performing my usual activities, I have moderate pain or discomfort, and I am not anxious or depressed' (examples taken from [34, 35]). As can be seen, in each description several domains of body structures and functions, as well as activities and participations are mentioned. Also, the description is not necessarily specific to a disease. That means that one health state can be present in different diseases, and that one disease can be the cause of several different health states.

It is worth noticing that, usually, all questionnaires from the second wave mentioned above (published after the '70s), describe health states in more or less similar domains of body structures, functions, activities, and participations. That shows a certain level of agreement about the core domains of functioning. The most commonly included domains are mobility, pain and discomfort, self-care, and mood.

However, knowing the health state does not in itself allow us to obtain a numerical representation of an individual's level of health, nor allow us to compare different health states numerically. The way to derive values (i.e. weights) from each health state (i.e. the valuation process) is described below.

2.2.4. Health - Disability weights and the valuation process

Valuating health states demands several considerations from different fields of knowledge: psychometrics, economic theory and ethics [6].

Comparing health states within the same domain of functioning can be considered a simple task since it reflects a natural intuition: someone with mild pain is healthier than someone with severe pain; someone who can get upstairs is healthier than someone who cannot. However, comparing health states across different domains may be more challenging: who is healthier, someone with moderate pain, or someone who cannot get upstairs? Even more complex interactions between different domains of the functioning can occur in the same individual, and the results of the health valuation for a given health state do not necessarily need to correspond to the linear sum of the weights of the compromised domains.

There are several procedures described in the literature for valuing health states, many of them based on social preferences. Some also include uncertainty in the process of choosing preferences.

The simplest procedure for valuating health states is through a visual analogue scale (VAS), where responders are required to localise the position of different health states in a scale, usually anchored between 0 and 100, where 0 means death and 100 the 'best imaginable health'. However, the use of VAS to obtain health/disability weights raises some concerns,

since there have been reports of a lack of linearity in answers regarding physiological stimuli [6].

One of the best-known procedures for health valuation is the standard gamble (SG). In this, individuals are required to choose between living, for example, ten years in a certain health state or living with perfect health after receiving a procedure with a risk of death. Identifying the risk of death (as a percentage) that the individual is willing to face allows us to assign a valuation to their health state [36]. Exploring social preferences considering uncertainty on the outcome of death is a distinctive element of this method, meaning it can be linked with the expected utility theory, which in turn, would be adequate for the resource allocation process [13, 36]. However, some critics of the SG point out that the result of the valuation process includes not only social preferences about health states, but also the aversion of individuals to the risk, which can bias the valuation [37].

Time trade-off (TTO) is another procedure where the responders are required to choose between living, for example, ten years in a particular health state, or living with perfect health but for a shorter number of years. The selected number of years lived with perfect health is used as a measure of preference [37]. Some concerns about this procedure are the potential threshold effects and the time preference bias [38]. Despite these concerns, authors have signalled that this method is the only one coherent with the metric unit used in summary measures that apply temporal frames, such as years of life [37].

Another common procedure for health valuation is the person trade-off (PTO). In this procedure, responders are required to choose between a programme that prevents the death of, for example, 100 fully healthy individuals, or another programme that prevents the onset of a health problem associated with a certain health state in a larger number of healthy people. The number of people selected as equivalent to preventing 100 people dying is used as a measure of preference for that health state [39]. However, it has been pointed out that some distributive considerations (e.g. some individuals prefer to avoid any death before preventing diseases in many other individuals), may confound the results of this method [6, 40].

All these procedures have been used in health economics for calculating weights that feed some of the metrics relied on upon cost-utility analyses, such as Quality-Adjusted Life Years (QALY) [41, 42]. This metric, although not traditionally used to measure population health, is

often used to assess the effectiveness of health interventions in terms of utilities. The use of utilities, where social preferences are considered, is consistent with the thinking of objective utilitarianism from the point of view of political philosophy [43]. Consistently, several authors have defined health state valuations as a measure of utilities associated with health states [36, 44], describing the weights as forms of health state utilities [45].

Different from these procedures, are other approaches that attempt to conduct health state valuation on a direct metric of health. For example, the WHO initiative of the Global Burden of Disease Study, and the later initiative of the global burden of diseases conducted by the IHME, have calculated weights through at least two different procedures.

The first was implemented during the '90s and early years of the 2000s for the initial versions of the WHO Global Burden of Disease Study. A long list of diseases was selected and their health states were described extensively. Then, through a multi-method process of expert elicitation, the weights for each health state were collected[46]. Similar approaches were replicated in several countries[47, 48].

The second approach was implemented for the second wave of the Global Burden of Disease studies. Through the various types of surveys, responders were invited to select 'the healthier state' between two competing states of health. Then a paired comparison procedure was used to create a scale of weights. Additionally, on a shortlist of health states, a PTO procedure was implemented to complement the procedure and estimate final weights. The information was collected through a general population-based survey carried out in several countries, as well as an electronic survey for experts and health scientists [34, 49, 50].

It is worth noting that the paired comparison procedure demands that responders to define what health state is healthier, which is different from demanding a preference about two health states in terms of risk of dying, time or number of persons. The conceptual difference is large, although in practice seems subtle. Comparisons between different procedures generally show a good correlation between different procedures [6]. The use of multi-method approaches for the health valuation process has also been proposed [6].

Some authors have suggested criteria for assessing the adequacy of the different procedures of health valuation. Among these criteria, the most important would be that the result of the valuation allows a meaningful interpretation in terms of years of longevity [37, 40]. This

criterion ensures that the weights for each health state can be combined with years lost to death. By that standard, all the other procedures presented (VAS, SG, TTO, PTO, Paired Comparison, Multi-methods) are anchored in extreme values, meaning either fully healthy or dead.

There are several other issues about the health valuation process that are less important as a background for this thesis, but that can be addressed in future work related to measuring effective coverage. Some of these issues are related to the thresholds needed to consider a decrement or improvement in functioning as a decrement or improvement in the health status of an individual; what technical aids such as orthosis or prosthesis can be included when the capacity and performance are assessed; whether health status from different ages should be valued differently; and whether a discount rate should be applied to the time lived with a certain health state [13, 51].

To summarise, in this section I have argued for the convenience of the ICF-2 framework to define the concepts of health and disability. In addition, I have briefly reviewed different alternatives for health valuation, which will help us to understand the pros and cons of the approach that I will propose in the following sections.

In the next section, I will introduce a questionnaire designed under the ICF-2's conceptual model to describe health states, which will be used to calculate the effective coverage. I will examine its psychometric properties in the context of the database where it was applied, and also introduce the use of a latent variable analysis to obtain disability weights through an alternative approach to those presented in this introduction.

2.3. Methods

2.3.1 The Health State Description questionnaire from the World Health Survey (WHS-HSD).

Different instruments have been used to describe health states; however, few of them have been developed explicitly under the framework of the ICF-2, proposed by the WHO [26, 28]. One of the first instruments used to describe health states under this framework was the Health State Description (HSD) questionnaire, designed for the World Health Survey (WHS) initiative, a large survey carried out in 70 countries between 2002 and 2003 [31]. Although the results obtained through this questionnaire advanced the understanding of the variability of health states across different populations, the literature that makes use of this instrument is relatively scarce and only related to the WHS initiative [52-55].

The WHS initiative was launched by the WHO, and its main purpose was to allow worldwide comparisons of health outcomes using valid, reliable and comparable methods [31]; [http://www.who.int/healthinfo/survey/en/ consulted in September 2018]. Among the instruments included in the survey, the HSD-WHS together with the Health State Valuation questionnaire were designed specifically to assess and compare the health of populations across the globe [13]. The instrument consists of 18 items, exploring overall health and eight different domains of functioning (mobility, self-care, pain and discomfort, cognition, interpersonal activities, vision, sleep and energy, and affect; see the questionnaire on Table S2.1 from the supplementary material). Each item enquires about the level of difficulty or any problems in performing different activities during the last 30 days. Responses to each item are structured using 5 Likert categories, ranging from 'none' to 'extreme or cannot do'. The HSD-WHS was developed by a panel of experts summoned by the WHO, and a pilot version was tested in the Multi-Country Survey initiative during 1999-2000 [13].

2.3.2. Data

To develop a procedure for estimating effective coverage, I needed a database where I could test each method associated with the metric. The required features for that database were: 1) the inclusion of a questionnaire that describe health states, underpinned in the framework of the ICF-2; 2) the presence of items and anthropometric or biologic measurements that identify people with several health conditions (i.e. with the need of a healthcare

intervention); 3) items that allow the identification of people who have access to those healthcare interventions; 4) a relatively large sample size to avoid underpowered statistical analysis; 5) high heterogeneity between responders in terms of health states, health conditions, access to healthcare interventions, and other covariables such as age, sex and markers of socioeconomic status; and 6) ideally a representative sample of a whole country. The last criterion was crucial since the estimator of effective coverage is thought to evaluate the performance of health systems at the country level or a comparably large catchment area. Each of these requirements will be introduced progressively throughout the subsequent chapters of this thesis.

I chose the database of a National Health Survey from Chile, which fulfils all these criteria and has the advantage of being very well known by the author, who also participated in its design. Moreover, that database is known by the first supervisor, who has already published other papers using this data source [56]. Using a familiar database simplified the process of exploring and developing the new metric.

Chile has carried out health surveys with nationally representative samples every six years since 2003. I used the second National Health Survey (Ch-NHS), undertaken between October 2009 and April 2010. The sample was drawn using a probabilistic, multi-stage and stratified design among people aged 15 or older. Eligible individuals were randomly chosen and interviewed at their homes. Pregnant women and institutionalised people (prison inmates, hospital patients) were excluded. A total of 5,412 interviews were completed (response rate 75%). Trained Interviewers followed a strict protocol and quality controls were implemented. Interviewees provided written informed consent. In cases of people younger than 18 years, assent was required along with written informed consent from their caregivers. The protocol was evaluated and accepted by the Ethics Committee of the Pontificia Universidad Católica de Chile. More information is available elsewhere [57]. Because Chile did not participate in the WHS, the Mexican Spanish version of the HSD-WHS questionnaire was used. No change in the language was required. To date, the psychometric properties of the Chilean version of the HSD-WHS have never been evaluated.

2.3.4. Covariates

In order to test the discriminant validity of the HSD-WHS, three chronic health conditions included in the Chilean National Health Survey were chosen, due to their high prevalence and local relevance: depressive episodes, hypertension and diabetes.

People with a depressive episode during the last 12 months were identified using the CIDI-Short Form questionnaire [58, 59], applying DSM-IV criteria. Hypertension was acknowledged when systolic or diastolic blood pressure exceeded 140 or 90 mmHg respectively after five minutes of rest, or in cases with normal blood pressure but a reported diagnosis along with current treatment for hypertension (lifestyle and/or drug). Similarly, diabetes cases were detected through a single fasting glycaemia of 125 mg/dl or higher, or normal glycaemia but with a report of diagnosed diabetes and current (drug) treatment for diabetes. Diabetes cases diagnosed during pregnancy were excluded.

Additionally, for descriptive purposes, age, sex and educational variables were also used in the analysis. Education background was categorised into three levels: ≤ 8 years, 9-12 years, and >12 years of formal education, respectively corresponding to primary, secondary, and high school or higher in Chile.

2.3.5. Statistical analyses

Descriptive characteristics of the sample in terms of age, sex, educational level and main morbidities were calculated.

To explore the dimensionality of the HSD-WHS, an exploratory factor analysis (EFA) using a polychoric correlation matrix from the answers and the weighted least square estimator [60] was conducted on a randomly selected half of the sample (n=2,646).

The EFA is commonly used to explore the relationship between observed variables and assumed latent variables (also called factors or common factors), which by definition are not observed [61]. In the HSD-WHS, the first-order of latent variables corresponds to each domain of health – in total nine domains – while the observed variables are the answer to

the 18 items of the questionnaire. In the EFA, the relationship between factors and observed variables is given by:

$$x = \Lambda f + u$$

where *x* is the set of observed variables, *f* is a vector of the (unobserved) factors, and Λ is the matrix of factor loadings (λ), which are coefficients that relate *x* with *f*, and are equivalent to the regression coefficients from a linear regression model. The parameter *u* corresponds to the vector of residuals from each equation, and it is assumed that they are uncorrelated with each other or with the factors. For convenience, the factors are standardised assuming a mean zero and a standard deviation of one. At least to begin with, factors are assumed to be uncorrelated, allowing us to interpret the factor loadings as the correlations of factors and observed variables. In this way, the variance of each observed variable can be split into two components: in the unique variance, that is the variance of the residuals (Ψ), and in the communality, that is the variance shared with other variables via the factors (the sum of the squared λ of each factor for an observed variable). Consequently, the population covariance matrix (Σ) of the observed variables under the EFA assumptions has the form:

$$\Sigma = \Lambda\Lambda' + \Psi \qquad \text{[Eq2.1]}$$

The aim of the EFA is to obtain the matrix of factor loadings that allows us to reproduce Σ. Minimum likelihood, ordinary least square, or weighted ordinary least square procedures, among others, can be implemented to find Λ [61]. I chose the weighted ordinary least square procedure because it has been suggested as more robust in cases where data is categorical [62].

However, there are several solutions for Λ that can satisfy the equation [Eq2.1], and therefore it is necessary to add additional constraints. One alternative would be allowing the first latent variable to account for the largest part of the communality, and then fitting the factor loadings for the second latent variable which could account for all the communality unrepresented by the first latent variable, and so forth. Another alternative is to 'rotate' the matrix through a matrix multiplication in $\Lambda^*=\Lambda M$, where *M* is a particular orthogonal matrix that keeps the underlying mathematical properties of the model. This alternative allows for higher interpretability of factor loadings in some cases [61].

In this study, I chose an Oblimin rotation [63] because it allows correlation between factors, which is in accordance with the conceptual framework of the questionnaire [13]. It is worth remembering that the HSD-WHS questionnaire has two items per domain of functioning (i.e. factors or latent variables), and it is expected that these domains are correlated since they are the expression of a second-order factor (or latent variable), which is the health.

Since the EFA must be performed on a preestablished number of factors, to select this number I carried out a parallel analysis on the polychoric matrix of correlations of items. In the parallel analysis, the magnitude of the communal variance from factors (or eigenvalue) obtained using the correlation matrix from the data is compared with the communal variance from factors obtained from random data. The comparison is repeated using a different number of factors. The point when the communal variance obtained from the data lies under the communal variance obtained through random data shows the number of factors that should be used in the EFA [64].

A polychoric correlation matrix (i.e. a specific correlation matrix for categorical variables) was used for the parallel analysis and EFA, since the responses from the items are categorical.

After the EFA, a confirmatory factor analysis (CFA) was conducted upon the other half of the split sample (N=2,647), which allows us to corroborate the dimensionality suggested by the EFA [65].

The CFA is a specific type of structural equation model (SEM), and as with any SEM, the CFA has three elements: random variables, structural coefficients and sometimes non-random variables. Random variables can be of three types: observed (or measured), unobserved (or latent) and disturbance (error variables). The measured and latent variables can also be categorised as endogenous and exogenous, depending on whether or not they are determined by variables within the model [66]. The general expression of an SEM encompasses specifications for latent variables and observed variables. The latent variable specification is:

$$\eta = B\eta + \Gamma\xi + \zeta$$

where η is the set of endogenous latent variables, ξ is the set of latent exogenous variables, ζ is the set of errors from the equations, B is a coefficient matrix of the endogenous variables, and Γ is a coefficient matrix of the exogenous variables. In this model it is assumed that errors have a mean of zero and are uncorrelated with ξ.

In turn, the specification for the observed variables is:

$$y = \Lambda_y \eta + \varepsilon$$
$$x = \Lambda_x \xi + \delta$$

where *y* represents the observed endogenous variables and *x* the observed exogenous variables, while Λ_y and Λ_x are the matrix of the coefficients that shows the relationship of *y* to η and *x* to ξ. The error of measurement is represented by ε and δ. As in the latent model, errors are assumed with a mean of zero and uncorrelated with η and ξ respectively.

Since I am exploring a second-order structure, where the second-order latent variable of health determines the first-order domains of health, which then determines the observed data, the aforementioned specifications can be simplified:

$$\eta = \Gamma\xi + \zeta$$
$$y = \Lambda_y \eta + \varepsilon$$

The expression related to *x* was dropped because there are no exogenous observed variables, while $B\eta$ was prescinded since there is only one second-order latent variable, and the first-order variables are not influencing each other [66]. In this specification, Λ_y is the matrix containing the 18 first-order coefficients, while the matrix Γ contains the nine second-order coefficients. Additionally, it can be specified Φ as the variance matrix of the second-order latent variable (ξ), and Ψ as the covariance matrix of the first-order latent variables (η). For Ψ we assumed a diagonal matrix, which represents the assumption that the first-order latent variables are uncorrelated to each other, except in the correlation accounted for by the second-order latent variable.

As for EFA, the parameters of the CFA are estimated through the basic assumption:

$$\Sigma = \Sigma(\theta)$$

where Σ is the population covariance matrix of observed variables and $\Sigma(\theta)$ is the covariance matrix written as a function of the free model parameters. A maximum likelihood procedure was performed to estimate such parameters.

I tested three models in CFA: a) a unidimensional model, b) a second-order model according to the structure suggested by the EFA, and c) a second-order model according to the structure suggested by the theoretical framework of the questionnaire [13]. The unidimensional model is the most parsimonious model, which assumes that all observed variables are correlated with only one first-order latent variable, health.

The goodness of fit for each CFA was evaluated using measures of relative and absolute fit, namely the relative χ^2 , the Tucker-Lewis index (TLI), the comparative fit index (CFI), the Goodness of Fit index (GFI), the root mean square error of approximation (RMSEA), and the standardised root mean square residual (SRMR) [67, 68]. Cut-off criteria for an adequate goodness of fit for these measures are relative $\chi^2 \leq 5$; TLI, CFI and GFI ≥ 0.95; RMSE ≤ 0.06; and SRMR ≤ 0.08. A detailed description of these measures is available elsewhere [67].

To study the discriminant validity (i.e. the evidence that points out if a latent variable differs according to other constructs in agreement with a theoretical framework [69]), the estimated value of the second-order latent disability variable was calculated for each participant of the survey through the Empirical Bayes Modal approach, also known as the regression method (n= 5,293). For interpretability reasons, the score was scaled from 0 to 100, where 0 is equivalent to having answered 'no difficulty or problem' in the 18 items of the questionnaire, while 100 means 'extreme difficulty or cannot do' on all the items. Since the items of the HDS-WHS do not allow a clear distinction between capacity and performance, I assumed that the health score = 100 - disability score [13] (see chapter one). I chose disability as a preferable expression of the results for interpretability reasons. To show whether the latent variable of disability behaved as expected, I compared its mean across people with and without chronic diseases, using lineal regression models, adjusted for age, gender and education. Later in the chapter, I will argue that the regression coefficients of these regression models can also be used as disability weights.

The reliability (internal consistency, or the coherence between the answers from different items) of the HSD-WHS was evaluated using Cronbach's alpha within each factor [70]. Values higher than 0.7 are usually considered acceptable [71]. Alpha is an expression of the correlation between two random sample of items from a universe of items like those in the questionnaire and is given by:

$$\rho_T = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n} \sigma_i^2}{\sigma_X^2}\right)$$

where ρ_T is the Cronbach's alpha, n is the number of items, σ_i^2 is the variance of the item *i*, and σ_X^2 is the variance of the sum of the scores from each item. This indicator of reliability can be interpreted as the inverse of the fraction of the variance of the total scores that follow from random error [72].

Regression models were conducted using the weights of the survey corresponding to the inverse of the probability of selection by each participant. A sensitivity analysis of the CFA using sample weights but assuming continuous variables was also performed. All analyses were performed using the statistical software R version 3.1.1, and using the packages *lavaan [73]*, *lavaan survey*, *polychor*, *psych*, *sem*, *semPlot* and *survey*.

2.4. Results

From the 5,412 people interviewed, 119 (2.2%) had missing answers in at least one item of the HSD-WHS and were excluded from the analysis, leaving 5,293 valid interviews. Half of the participants were women and half were aged between 30 and 59 years. People younger than 30 represented roughly one-fifth of the sample, though after using the survey's sampling weights, this group accounted for 30.1% of the Chilean population older than 14 years. Nearly three-quarters of the participants reported more than eight years of education. Among the explored diseases, hypertension was the most prevalent condition (27.3% [25.2 – 29.4]), followed by a depressive episode (17.7% [15.8-19.6]), and then diabetes (6.8% [5.7-8.0]) (see Table 2.1).

Table 2.1 Description of the sample, Chilean National Health Survey 2009-2010 (n=5,293)

	N	%	CI95%
Sex			
female	3,137	51.3	[48.9 – 53.6]
Age			
15-29	1,190	30.1	[28.0 – 32.3]
30-59	2,716	51.6	[49.3 – 53.9]
60+	1,387	18.3	[16.7 – 19.9]
Education†			
≤8 years	1,415	19.0	[17.4 – 20.6]
9-12 years	2,884	56.8	[54.5 – 59.1]
>12 years	986	24,2	[22.0 – 26.4]
Morbidity*			
Depressive episode	743	17.7	[15.8 – 19.6]
Hypertension	1,523	27.3	[25.2 – 29.4]
Diabetes	379	6.8	[5.7 – 8.0]

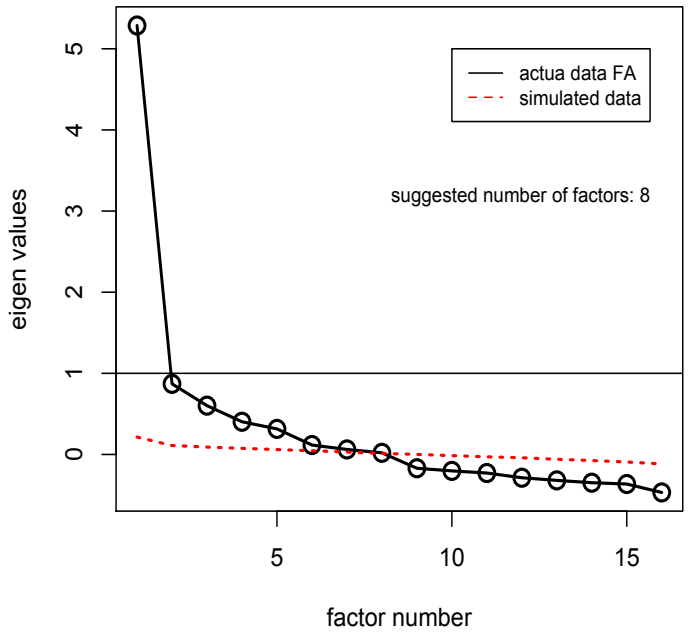
† prevalence of education was estimated over 5,285 cases without missing data
** prevalence of morbidities was estimated over 4,600 cases without missing data in measurement of blood pressure, fasting glycaemia or missing data in the HSD-WHS questionnaire*

Overall, the Keiser-Meyer-Olkin measure of sampling adequacy was 0.894, and higher than 0.8 for all items of the HSD-WHS questionnaire, while the Barret spherical measure was significant ($\chi 2=67.671,7$, df=153, p-value <0.001). These indices indicate that the data was suitable for a factor analysis. In total, 92.0% of polychoric inter-item correlations were equal to or higher than 0.3, and 92.3% of values were between 0.2 and 0.8 (see Table S2.2 and

Figure S2.1 from supplementary material). This is consistent with the presence of at least one underlying common factor.

Parallel analysis, for choosing the number of factors (or latent variables) that will be explored in the EFA, suggested eight factors. See Figure 2.2.

Figure 2.2. Parallel analysis for choosing the number of factors for the Health Sate Description questionnaire from the World Health Surveys initiative (HSD-WHS), Chilean National Health Survey 2009-2010 (N= 2,646).



The result of the EFA in terms of loading factors is shown in Table 2.1. This solution combined the original items from the mobility domain with those from overall health in a single factor or latent variable. Also, the factor for affect included an item from overall health, though with a small loading. Other items were grouped according to the original theoretical structure of the questionnaire [13]. In all cases, the uniqueness of items was under 50%. The total variance captured by the eight factors was 74%.

Table 2.2. Exploratory Factor Analysis: Factor loadings of each item with the eight subdomains of the Health State Description questionnaire of the World Health Survey (HSD-WHS), Chilean National Health Survey 2009/2010.

Item	Overall Health & Mobility	Self-Care	Pain & Discomfort	Cognition	Interpersonal activities	Vision	Sleep & Energy	Affect	Uniqueness
1	**0.43**	-	-	-	-	-	-	*0.27*	48%
2	**0.61**	-	-	-	-	-	-	-	27%
3	**0.76**	-	-	-	-	-	-	-	12%
4	**0.49**	-	-	-	-	-	-	-	53%
5	-	**0.74**	-	-	-	-	-	-	14%
6	-	**0.99**	-	-	-	-	-	-	0%
7	-	-	**0.74**	-	-	-	-	-	27%
8	-	-	**1.01**	-	-	-	-	-	1%
9	-	-	-	**1.00**	-	-	-	-	0%
10	-	-	-	**0.57**	-	-	-	-	36%
11	-	-	-	-	**0.98**	-	-	-	1%
12	-	-	-	-	**0.50**	-	-	-	51%
13	-	-	-	-	-	**0.81**	-	-	33%
14	-	-	-	-	-	**0.83**	-	-	31%
15	-	-	-	-	-	-	**0.78**	-	37%
16	-	-	-	-	-	-	**0.69**	-	37%
17	-	-	-	-	-	-	-	**0.82**	23%
18	-	-	-	-	-	-	-	**0.75**	30%

Factor loadings lower than 0.2 are omitted.
Larger factor loadings are marked with bold. Italic numbers show cross-loadings.
n 2,646/ method: weighted least square/ rotation: oblimin.

The CFA according to the theoretical model and the model suggested by the EFA showed an adequate goodness of fit. The unidimensional model showed worse parameters for the goodness of fit than the alternatives (see Table 2.3). Path diagrams of the models are presented in Figure 2.2, where it is observed that the latent variable of disability determines the values of the latent variables for each domain, which in turn determines the answer for each item.

Table 2.3. Model fit for confirmatory factor analysis for the Health State Description questionnaire of the World Health Survey (HSD-WHS), Chilean National Health Survey 2009/2010. (n=2,647).

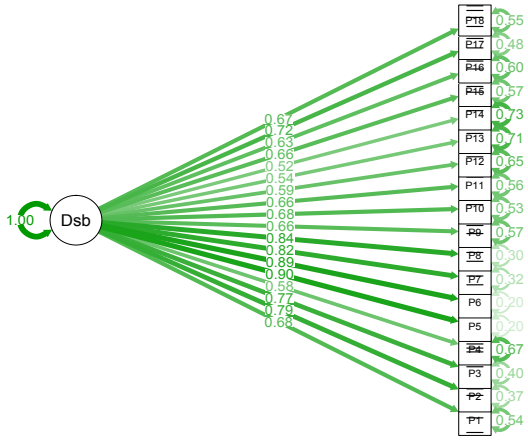
	Unidimentional model	Model suggested by EFA	Theoretical model
χ^2	7051.0	1571.1	1636.7
Df	135	127	126
p-value	<0.001	<0.001	<0.001
relative χ^2	52.2	12.4	13.0
TLI	0.94	0.99	0.99
CFI	0.95	0.99	0.99
GFI	0.96	0.99	0.99
RMSEA	0.139	0.066	0.067
RMSEA LCI	0.136	0.066	0.064
RMSA UCI	0.142	0.068	0.070
SRMR	0.114	0.064	0.064

EFA: Exploratory factor analysis / χ^2: chi-square / Df: degree of freedom / TLI: Tucker-Lewis index / CFI: comparative ftr index / GFI: goodness of fit index / RMSEA: root mean square error of approximation / SRMR: standardised root mean square residual / LCI: lower 90% confidence interval / UCI: upper 90% confidence interval

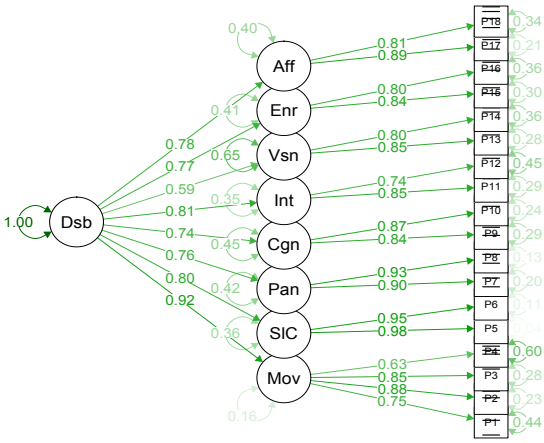
No statistical difference was observed in the goodness of fit between the theoretical model and that suggested by the EFA (χ^2 *diff=65.6, df diff=1, p-value=1.00*) [68]; thus, the former was chosen as the basal model. The correlation matrix between factors and additional information is shown in the supplementary material: Tables S2.3 - S2.5 and Figures S2.2 - S2.3. The overall goodness of fit of the basal model in the total sample was: *X^2= 2558.9, df=126, p-value<0.001, relative χ^2 = 20.3, TLI= 0.99, CFI= 0.99, GFI= 0.99, RMSA= 0.060 [90%CI 0.058 – 0.062], SRMR= 0.056.* A description of the scaled factor scores (latent variables) from the basal model, calculated using the whole sample, is presented in Table 2.3.

Figure 2.3. Path-diagram of the Health State Description of the World Health Survey (HSD-WHS) in the Chilean National Health Survey 2009/2010. (n=2,646).

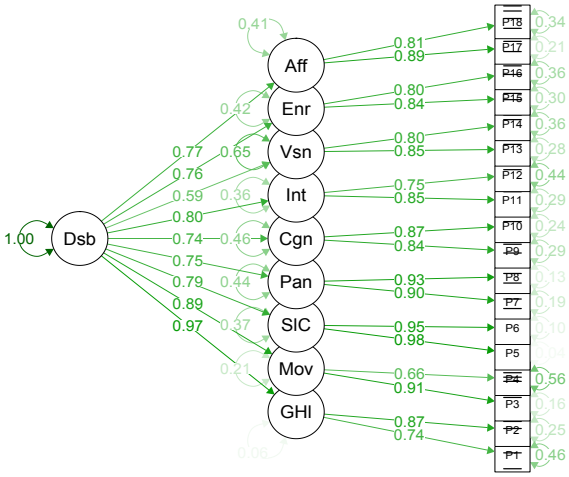
A. Unidimensional model by EFA



B. Model according to the structure suggested



C. Model according to the structure suggested by the theoretical framework



EFA: exploratory factor analysis

Dsb: disability/ Mov: mobility domain/ SLC: self-care domain/ Pan: pain and discomfort domain/ Cgn: cognition domain/ Int: interpersonal activities domain/ Vsn: vision domain/ Enr: sleep and energy domain/ Aff: affect domain/ GHI: Overall Health/ P1-P18: represent the 18 items of the questionnaire.

Values on the straight arrows are the standardised coefficients/ values on the curved arrows are the unexplained variance of the variables.

Table 2.3. Reliability, intra-class correlation of the Health State Description questionnaire of the World Health Survey (HSD-WHS), and description of factors scores (latent variables), Chilean National Health Survey 2009/2010. (n=5,293).

	std.alpha	mean	sd	median	Q1	Q3	Minimum	Maximum
Overall Health	0.70	32.2	14.9	31.1	21.3	42.1	0.0	100.0
Mobility	0.65	32.3	15.7	30.9	21.2	41.5	0.0	100.0
Self Care	0.88	32.2	15.9	30.8	21.0	41.3	0.0	100.0
Pain & Discomfort	0.88	32.2	19.0	30.2	16.7	45.5	0.0	100.0
Cognition	0.77	30.0	17.2	28.4	15.7	41.2	0.0	100.0
Interpersonal	0.67	30.8	15.7	29.1	19.1	40.6	0.0	100.0
Vision	0.71	27.1	17.6	21.4	13.4	39.3	0.0	100.0
Sleep & Energy	0.72	31.9	17.1	30.6	18.4	43.8	0.0	100.0
Affects	0.79	33.3	18.0	32.0	19.2	45.5	0.0	100.0
Disability	-	32.0	14.7	31.2	21.4	41.6	0.0	100.0
Whole instrument	0.90	-	-	-	-	-	-	-

stda.alpha: standardised alpha/ sd: standard deviation / Q1: quantile 25% / Q3: quantile 75%

The HSD-WHS questionnaire showed a good overall reliability (standardised Cronbach alpha = 0.90). Mobility and interpersonal activities were the domains with lowest standardised Cronbach alpha, although higher than 0.6 in both cases (see Table 2.3). Median values of the factor scores (latent variables) were similar to the means. The only factor that showed a more skewed distribution was 'vision'.

The mean disability scores for each disease, as well as for sociodemographic variables, are presented in Table 2.4. Disability was higher among women and older people, while lower for those with higher levels of education. Results are consistent with those presented using bivariate and fully adjusted regression models. Among diseases, in the multivariate model depression showed the highest disability score of 13.6 [12.1 - 15.2], on a scale between 0 and 100, followed by diabetes and hypertension. In the discussion, I will argue that these disability scores can also be interpreted as disability weights. These results were obtained using the sample with complete data for all covariables (n=4.600)

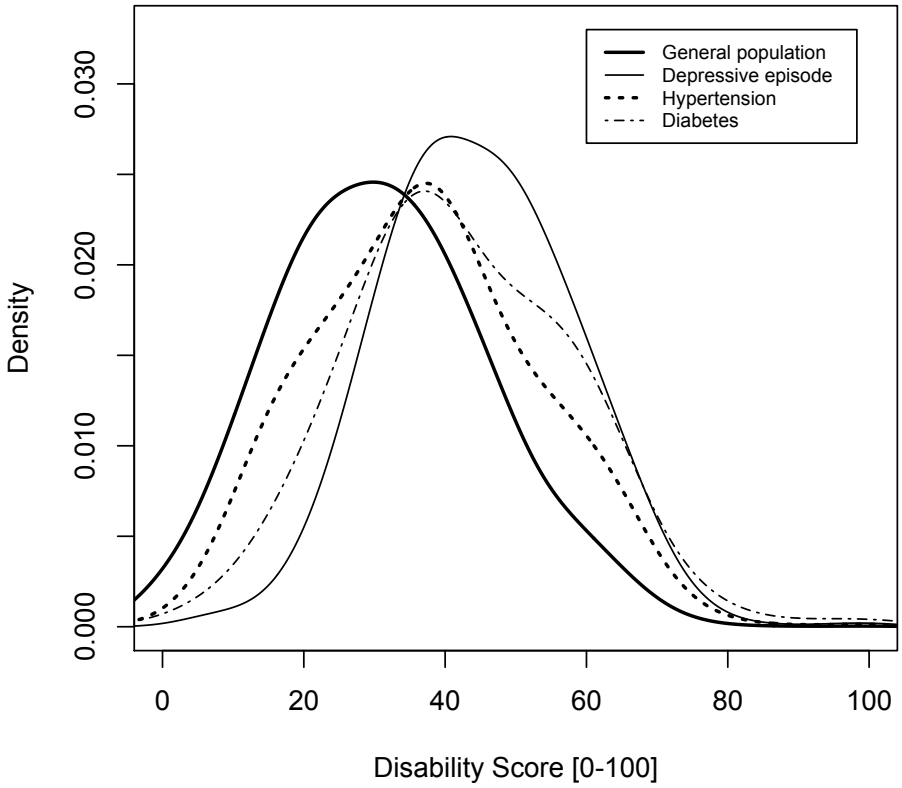
Table 2.4. Means and coefficients of regression models for the latent variable of disability according to sociodemographic variables and morbidity, Chilean National Health Survey 2009/2010 (n=4,600).

	Mean of Disability score					Coefficients from regression model for Disability			
	mean	**CI 95%**	**mean**	**CI 95%**		**Coefficient**	**CI 95%**	**Coefficient**	**CI 95%**
	Women		Men			Bivariate		Multivariate	
					Intercept	-	-	21.5	[18.8 to 24.2]
Sex	35.6	[34.7 - 36.4]	28.1	[27.0 - 29.3]	Sex (women)	7.4	[6.0 to 8.9]	4.6	[3.4 to 5.9]
					Age (each 10 years)	2.8	[2.5 to 3.2]	1.9	[1.5 to 2.3]
	Whole sample		-						
Age categoric									
15 – 29 years	26.6	[25.4 - 27.8]	-	-		-	-	-	-
30 – 59 years	32.9	[31.8 - 33.9]	-	-		-	-	-	-
>59 years	38.9	[37.2 - 40.5]	-	-		-	-	-	-
Education			-		Education				
<8 years	39.4	[37.7 - 41.1]	-	-	*<8 years*	0.0	-	0.0	-
8-12 years	31.3	[30.4 - 32.2]	-	-	*8-12 years*	-8.1	[-10.0 to -6.2]	-3.1	[-4.9 to -1.4]
>12 years	28.0	[26.4 - 29.5]	-	-	*>12 years*	-11.04	[-13.7 to -9.2]	-5.2	[-7.4 to -3.1]
	With the condition		Without the condition						
Depressive episode	44.7	[43.1 - 46.3]	29.3	[28.5 - 30.0]	Depressive episode	15.4	[13.6 to 17.1]	13.6	[12.1 to 15.2]
Hypertension	37.6	[36.1 - 39.0]	29.9	[29.1 - 30.7]	Hypertension	7.7	[6.0 to 9.3]	1.6	[0.0 to 3.3]
Diabetes	42.3	[39.8 - 44.8]	31.2	[30.5 – 32.0]	Diabetes	11.0	[8.4 to 13.7]	5.0	[2.5 to 7.4]

CI: confidence interval

To evaluate the discriminant validity of the scale, in Figure 2.4 I present the Kernel distributions of the score of the disability latent variable for the general population and people with specific diseases. The distribution of cases with hypertension, diabetes, or a depressive episode clearly shifted to the right, meaning higher levels of disability.

Figure 2.4. Distribution of the latent variable of disability in the general population and among those with three health conditions, Chilean National Health Survey 2009/2010. (n=4,600)



A sensitivity analysis was performed, repeating the EFA and CFA and the reliability appraisal described above, using the 4,600 cases with complete data. No substantial changes in the results were observed (see supplementary material: Tables S2.10 - S2.12, and Figures S2.8 – S2.9). The results of a second sensitivity analysis for CFA using the sample weights, although assuming continuous variables, are presented in the supplementary material (see Tables S2.6 - S2.9, and Figures S2.4 - 2.7). The resultant factor scores (latent variables) were skewed towards the left, affecting the means but keeping the disability scores almost unchanged.

2.5. Discussion of results

To the best of my knowledge, this is the first psychometric evaluation of the HSD-WHS questionnaire in a population not included in the original WHO-Household Surveys Programme. In Chile, the HSD-WHS questionnaire showed good reliability and an adequate construct validity. The latent disability variable based on this questionnaire also showed a good discriminant validity, in accordance with different diseases and socioeconomic variables.

The HSD-WHS questionnaire was designed and built using a rigorous protocol and a robust conceptual framework. During the WHS initiative, it was tested in numerous countries across the world [13, 31, 74]. The results provided by the HSD-WHS during the WHS have also been useful for developing more recent disability questionnaires, such as the last version of the WHO Disability Assessment Schedule (the WHODAS 2.0) [75], widely used in the current scientific literature [76].

One of the strengths of the HSD-WHS questionnaire is its consistency, with a clear conceptual framework that provides explicit definitions for health and disability. The adherence by the HSD-WHS to the ICF-2 framework is an important difference in respect to other generic instruments widely used to assess health states and/or disability. For example, in comparison with the SF-36, the latter claims to measure 'functional status' and wellbeing, and following its technical documentation, gathers the eight 'health concepts' most frequently included in widely used surveys [21]. Furthermore, the SF-36 questionnaire does not use a unique statement for items, nor uses homogeneous response categories, and includes different standards of comparison between items. These elements clearly distance the SF-36 from the concepts of health and disability proposed by the ICF-2.

Another widely used instrument to assess health states is the EuroQol EQ-5D questionnaire mentioned previously [77]. Its final purpose is not only to identify different health states, but also to provide social preferences or utilities for each one of them (i.e. health state utilities). In contrast with the HSD-WHS questionnaire, the EQ-5D explores only five domains of functioning, using three Likert categories as answers, which allows it to generate 243 (3^5) different health states. On the other hand, the HSD-WHS explores nine domains of functioning through 18 items, using five Likert categories as answers. This allows the HSD-

WHS to potentially generate 5^{18} health states, which means more than one billion alternatives and which clearly reflects the continuous nature of the latent constructs of health and disability. Moreover, some important domains of functioning are not included in the EQ-5D: for example, those related to sense organs [78]. Conversely, the HSD-WHS includes two items to assess vision, and also gives more relevance than the EQ-5D to domains such as sleep, energy, depression and anxiety. In addition, EQ-5D has been criticised for using three Likert categories, which could produce ceiling effects and a lack of sensitivity to changes in health[78]. The latest versions of the EQ-5D have tried to overcome this limitation by including five Likert categories, keeping constant the five dimensions (EQ-5D-5L). However, this generates 3,125 (5^5) health states that need to be assessed under TTO or another procedure, making more complex the health valuation process.

However, the most important difference between HSD-WHS and the other questionnaires used for describing health states is not how they assess each state, but how they are used to generate weights for such health states. In the background section of this chapter, I mentioned that the procedures for valuing health states can be broadly classified into two different categories: those that rely on social preferences such as VAS, SG, TTO and PTO, which are commonly used in cost-utility analysis; and those procedures that pretend to assess directly the constructs of health through procedures of paired comparisons, as in the case of DALYs calculation in the Global Burden of Disease initiatives. It is worth noting that some authors have claimed the paired comparison carried out by the IHME in the burden of disease studies is just another way to collect social preferences about health states [79].

In this study, to ascertain the discriminant validity I proposed using a latent variable approach, based on an SEM for valuating health states, which is a completely different procedure to all the methods described above. The weights of the health states generated through the latent variables approach are calculated using the parameters estimated for the SEM as CFA. Giving a specific combination of answers for the 18 items of the HSD-WHS and applying them to the SEM's parameters, it is possible to calculate specific values (i.e. disability weights) for each of the 5^{18} different health states that the questionnaire can explore. However, health state weights alone are of little use. Usually, they have to be associated with a health condition, via the selection of a set of health states for a specific disease, weighting them according to their relative frequency.

For example, using the National Health Survey database I can select all the individuals with a depressive episode, extract their associated health states, and calculate the predicted values of the latent variable of disability for each one of those health states (i.e. disability weights). Then I can multiply them by their frequency and sum them. In this way, I would obtain an expected value of the disability for individuals with that disorder. Another alternative is to ask a group of experts or patients with the diagnosis of depression to choose the health states that best describe the condition, according to the items of the HDS-WHS. Then I can use the disability weights, already calculated from the national representative database, for each health state and calculate the expected value of disability. This method (i.e. through expert or patient elicitation) is a common way that the health state values are developed to produce summary measures such as QALYs or DALYs for specific conditions that, for instance, are not included in the database from which the weights are derived [41].

In the method proposed I followed a different approach to calculate summary measures of disability for health conditions. I used a regression model to estimate the expected value of disability associated with each disease. For example, in the bivariate analysis it is observed that depressive episodes had a regression coefficient of 15.4 [13.6 to 17.1], which means that people with a depressive episode have on average a 15.4 higher score for disability than people without a depressive episode. This is consistent with the raw difference between the means of the disability score that is also reported in people with (44.7) and without (29.3) that condition. However, I recognise that at least part of the disability reported among people with or without a depressive episode comes from other factors – for example comorbidities – which are not necessarily equally distributed between both groups of people. For that reason, I also report the multivariate results, where regression coefficients can be interpreted as the disability accounted by each variable adjusted by the others. In the multivariate regression model, the disability score for depression diminishes to 13.6 [12.1 - 15.2]. I propose using that score as the health state value for people with a depressive episode for the target population where the data is coming from, in this case, Chilean adults from 2009-2010. I am assuming that all relevant confounders for the relationship between depressive episodes and disability were considered, and the cases of depressive episodes were correctly identified. It is important to remember that the scale of the health valuation, in this case, ranges between absence (0) and extreme difficulty (100) performing activities in all items of the HSD-WHS questionnaire.

To use this procedure, based on regression models, for calculating health state values for health conditions seems convenient because of its simplicity and flexibility. This approach allows us to easily adjust the disability weights by multimorbidity and even calculate them considering interaction terms between diseases. The use of this approach is also convenient for calculating attributable fractions, in order to estimate the specific contribution of diseases to the total population disability [80]. This is a crucial procedure for estimating effective coverage and is further developed in chapter three of this thesis.

Just to provide a meaningful comparison, following the example of depression, the IHME in its study of 2015 calculated a disability weight for a mild major depressive disorder of 0.145 [0.099 - 0.209] on a scale from full health (0) to death (1) [34], which is not far from the value I calculated for Chile using the procedure described above: 13.6. If we use health state values (or health state utilities, where social preferences are included) obtained from calculating QALYs, some reports suggest scores of 0.700 [0.670 – 0.730] for remission and 0.570 [0.540 - 0.610] for minor depression [81], in a scale from death (0) to full health (1). The latter values can be compared with the means of the disability score I obtained for people without a depressive episode (29.3, where its complement is 100-29.3 = 70.7) and with a depressive episode (44.7, where its complement is 100-44.7 = 59.3), which shows that my results are within the range of other valuations. In chapter four, I provide disability weights for different levels of severity of depressive episodes using the Chilean data. These comparisons are only shown to support that the obtained results are plausible. The small differences in numbers can be attributable to difference in the methodological approach and because the results obtained in this Thesis are based directly on general Chilean population.

Another interesting element to discuss is the proximity between items from questionnaires used to identify people with depression with those used for measuring disability (i.e., activities and participation). This is evident when we compare the 17th item of the HSD-WHS: “how much of a problem did you have with feeling sad, low or depressed”, and the first item of the CIDI-SF: “have you had two weeks or more in a row when you felt sad, low, or depressed?”. This is a common feature for almost all psychiatric measures, where items used to capture the presence of symptoms are closely related to those measuring functioning. The similarity between items can be seen as problematic because we tend to think that there is a clear distinction between mental disorders and the disability produced by them.

However, this happens more often when we compare isolated items, but not when we compare the whole questionnaires and, more importantly, how these items are used to build the latent construct that underlies them. In the case of depression, for example, the first item of the CIDI-SF questionnaire is combined with loss or gain of body weight and suicidal ideation, amongst others, and items are classified in major and minor criteria according to clear rules from the DSM-IV. Meanwhile, item 17 of the HSD-WHS tends to cluster with other items asking about problems such as seeing, walking or pain following coefficients through a structural equation model that generates a latent variable conceptually linked with disability. In other words, the "budling blocks" can be similar, but the resulting buildings are different.

One limitation of my approach is that the scale I produced is anchored in extreme values that are not meaningful in terms of longevity – a requirement for health metrics that are combined with a component of mortality and are useful for allocating resources [37]. This limitation can be resolved if the scale of disability is anchored in the range between full health and death, and not, as now, between absence and extreme difficulty in performing activities. The absence of difficulty can be assimilated to full health, but to find the position of extreme difficulty in respect to death is more complex. It is worth noting that, in the case of the DALYs, health states worse than death are not allowed. However, in the case of health state utilities obtained through the EQ-5D used in QALYs, it is common to find that roughly a third of all values are worse than death, reporting health state utilities lower than zero [82]. However, for the purposes of this thesis, the re-scaling is not needed, because I will focus the effective coverage estimator on avoided disability, leaving the integration of mortality for future developments (see chapter six).

To value health states through a direct metric of health in opposition to social preferences has important theoretical consequences. One of them is that, through the latent variables method, I distance myself from the concept of utilities – we can call it the 'QALY approach' – which has been argued as the most adequate metric to allocate resources [37]. On the other hand, I am coming closer to the metrics that pretend to directly monitor the health of populations, which can be called the 'DALY approach' [5]. The convenience of using this approach to measure effective coverage will be discussed in the following chapters.

Another limitation of this study is the absence of anchoring vignettes that allow us to adjust the responses of individuals to social expectations concerned with their health status [83]. Consequently, it is assumed that disability or health are homogeneous within the population,

regardless of their cultural or socioeconomic position. The possibility of a response bias according to these features cannot be completely ruled out, nor can it be resolved by adjusting the regression models by those factors. However, this limitation can be overcome in future studies by implementing a variance analysis [66] or by Multiple Indicator Multiple Causes models [84], when the latent variables are being calculated.

In addition, I did not test the convergent or predictive validity of the HSD-WHS, neither did test-retest reliability. Although discriminative validity was evaluated using three common morbidities and the usual sociodemographic variables, a more in-depth analysis could have been performed using a greater number of morbidities and additional sociodemographic variables. The HSD-WHS questionnaire was implemented in the context of a National Health Survey and it did not include repeating measurements with subsamples, nor include measures representing the gold standard. Notwithstanding these limitations, the robust conceptual framework that backs up this instrument, and the current report of its psychometric properties, provides support for its use in future studies.

To summarise, in this chapter I have proved the adequacy of the HSD-WHS to describe health states and proposed a method based on latent variables and regression models to calculate health state values for diseases, all under the conceptual framework of the ICF-2.

The use of latent variables to measure disability through the HSD-WHS questionnaire also opens up a valuable opportunity to explore different approaches towards measuring the contribution of diseases, as well as social determinants, to disability.

2.6. Introduction to the next chapters

In this chapter, I defined what is understood by health and disability in the context of this thesis. I also discussed how to describe health states and, most importantly, how to value them. I showed that the psychometric properties of the questionnaire I will use to measure disability are adequate in the context of the Chilean population. Also, I proposed a specific approach to obtain summary health valuations for diseases.

In the next chapter, I will use these elements for calculating the fraction of disability at the individual and population level that is attributable to different diseases. The procedure to obtain such attributable fractions is a crucial component of the overall method to estimate effective coverage.

References of chapter two

1. Shengelia B, Tandon A, B. Adams O, Murray CJL: **Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy**. *Social Science & Medicine* 2005, **61**:97-109.
2. World Health Organization: **Tracking universal health coverage: first global monitoring report**. In.; 2015.
3. European Observatory on Health Systems and Policies: **Health system performance comparison. An agenda for policy, information and research**: Open University Press; 2013.
4. Murray CJL, Frenk J: **A framework for assessing the performance of health systems**. *Bulletin of the World Health Organization* 2000, **78**(6):717-731.
5. Murray CJ, Salomon JA, Mathers C: **A critical examination of summary measures of population health**. *Bull World Health Organ* 2000, **78**(8):981-994.
6. Salomon JA, Murray CJ: **A multi-method approach to measuring health-state valuations**. *Health Econ* 2004, **13**(3):281-290.
7. Organization WH: **Constitution of the world health organization**. In: *Official Records of the World Health Organization.* Edited by Organization WH, vol. 2. New york: World Health Organization; 1946.
8. Chen MK, Bryant BE: **The measurement of health--a critical and selective overview**. *Int J Epidemiol* 1975, **4**(4):257-264.
9. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW: **Studies of Illness in the Aged: The Index of ADL: A Standardized Measure of Biological and Psychosocial Function**. *JAMA* 1963, **185**(12):914-919.
10. Abramson JH: **The cornell medical index as an epidemiological tool**. *Am J Public Health Nations Health* 1966, **56**(2):287-298.
11. Grogono AW, Woodgate DJ: **Index for measuring health**. *Lancet* 1971, **2**(7732):1024-1026.
12. Breslow L: **A quantitative approach to the World Health Organization definition of health: physical, mental and social well-being**. *Int J Epidemiol* 1972, **1**(4):347-355.
13. Salomon JA, Mathers CD, Chatterji S, Sadana R, Üstün TB, Murray CJL: **Quantifying Individual Levels of Heath: Definitions, Concepts, and MEasurement Issues**. In: *Health Systems Performance Assessment Debates, Methods and Empiricism.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 301-318.
14. Boorse C: **Health as a Theoretical Concept**. *Philosophy of Science* 1977, **44**(4):542-573.
15. Venkatapuram S: **Health, vital goals, and central human capabilities**. *Bioethics* 2013, **27**(5):271-279.
16. Edwards SD: **Nordenfelt's theory of disability**. *Theor Med Bioeth* 1998, **19**(1):89-100.

17. Organization WH: **International Classification of Impairments, Disabilities, and Handicaps. A manual of classification relating to the consequences of disease.** Geneva: World Health Organization; 1980.
18. Schuntermann MF: **The International Classification of Impairments, Disabilities and Handicaps (ICIDH)--results and problems**. *Int J Rehabil Res* 1996, **19**(1):1-11.
19. Pfeiffer D: **The ICIDH and the Need for its Revision**. *Disability & Society* 1998, **13**(4):503-523.
20. **EuroQol--a new facility for the measurement of health-related quality of life**. *Health Policy* 1990, **16**(3):199-208.
21. Ware JE, Sherbourne CD: **The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection**. *Medical Care* 1992, **30**(6):473-483.
22. Ustün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, Saxena S, von Korff M, Pull C: **Developing the World Health Organization Disability Assessment Schedule 2.0**. *Bull World Health Organ* 2010, **88**(11):815-823.
23. Federici S, Bracalenti M, Meloni F, Luciano JV: **World Health Organization disability assessment schedule 2.0: An international systematic review**. *Disabil Rehabil* 2017, **39**(23):2347-2380.
24. Hunt SM, McKenna SP, McEwen J, Williams J, Papp E: **The Nottingham Health Profile: subjective health status and medical consultations**. *Soc Sci Med A* 1981, **15**(3 Pt 1):221-229.
25. Parkerson GR, Jr., Broadhead WE, Tse CK: **The Duke Health Profile. A 17-item measure of health and dysfunction**. *Med Care* 1990, **28**(11):1056-1072.
26. Üstün TB, Chatterji S, Bickenbach J, Kostanjsek N, Schneider M: **The International Classification of Functioning, Disability and Health: a new tool for understanding disability and health**. *Disability and Rehabilitation* 2003, **25**(11-12):565-571.
27. Ayuso-Mateos JL, Nieto-Moreno M, Sánchez-Moreno J, Vázquez-Barquero JL: **Clasificación Internacional del Funcionamiento, la Discapacidad y la Salud (CIF): aplicabilidad y utilidad en la práctica clínica**. *Med Clin (Barc)* 2006, **126**(12):461-466.
28. World Health Organization: **International Classification of Functioning, Disability and Health (ICF)**: World Health Organization; 2001.
29. World Health Organization: **The World Health Report 2000. Health Systems: Improving Performance.** : World Health Organization; 2000.
30. Bank TIBfRaDTW: **Global Burden of Disease and Risk Factors**: Oxford University Press and The World Bank; 2006.
31. Üstün TB, Chatterji S, Mechbal A, Murray CJL, Groups WC: **The World Health Surveys**. In: *Health Systems Performance Assessment Debates, Methods and Empiricsm.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 797 - 808.
32. Nguyen T, Stewart D, Rosenbaum P, Baptiste S, Kraus de Camargo O, Gorter JW: **Using the ICF in transition research and practice? Lessons from a scoping review**. *Res Dev Disabil* 2018, **72**:225-239.

33. Maritz R, Aronsky D, Prodinger B: **The International Classification of Functioning, Disability and Health (ICF) in Electronic Health Records. A Systematic Literature Review**. *Appl Clin Inform* 2017, **8**(3):964-980.
34. Salomon JA, Haagsma JA, Davis A, de Noordhout CM, Polinder S, Havelaar AH, Cassini A, Devleesschauwer B, Kretzschmar M, Speybroeck N *et al*: **Disability weights for the Global Burden of Disease 2013 study**. *The Lancet Global health* 2015, **3**(11):e712-723.
35. Foundation ER: **EQ-5D-3L User Guide, 2018**. In., vol. 2018: EuroQol Research Foundation.
36. Torrance GW: **Social preferences for health states: An empirical evaluation of three measurement techniques**. *Socio-Economic Planning Sciences* 1976, **10**(3):129-136.
37. Richardson J: **Cost utility analysis: what should be measured?** *Social science & medicine (1982)* 1994, **39**(1):7-21.
38. Johannesson M, Pliskin JS, Weinstein MC: **A note on QALYs, time tradeoff, and discounting**. *Med Decis Making* 1994, **14**(2):188-193.
39. Nord E: **The person-trade-off approach to valuing health care programs**. *Med Decis Making* 1995, **15**(3):201-208.
40. Nord E: **Methods for quality adjustment of life years**. *Social science & medicine (1982)* 1992, **34**(5):559-569.
41. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, GW T: **Methods for the Economic Evaluation of Health Care Programmes**, 4th edn. Oxford, UK: Oxford University Press; 2015.
42. Weinstein MC, Torrance G, McGuire A: **QALYs: the basics**. *Value Health* 2009, **12 Suppl 1**:S5-9.
43. Roberts MJ, Hsiao W, Berman P, Reich MR: **Judging Health-Sector Performance: Ethical Theory**. In: *Getting Health Reform Right A Guide to improving Performance and Equity.* edn.: Oxford University Press; 2008.
44. Torrance GW: **Toward a utility theory foundation for health status index models**. *Health Serv Res* 1976, **11**(4):349-369.
45. Ara R, Wailoo A: **Using health state utility values in models exploring the cost-effectiveness of health technologies**. *Value Health* 2012, **15**(6):971-974.
46. Murray CJ, Lopez AD: **Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: global Burden of Disease Study**. *Lancet* 1997, **349**(9062):1347-1352.
47. Melse JM, Essink-Bot ML, Kramers PG, Hoeymans N: **A national burden of disease calculation: Dutch disability-adjusted life-years. Dutch Burden of Disease Group**. *Am J Public Health* 2000, **90**(8):1241-1247.
48. Group PH: **Victorian Burden of Disease Study. Mortality and morbidity in 2001.**: Public Health Group, Rural and Regional Health and Aged Care Services Division. Victorian Government Department of Human Services. Melbourne Victoria; 2005.
49. Haagsma JA, Maertens de Noordhout C, Polinder S, Vos T, Havelaar AH, Cassini A, Devleesschauwer B, Kretzschmar ME, Speybroeck N, Salomon JA: **Assessing disability weights based on the responses of 30,660 people from four European countries**. *Population health metrics* 2015, **13**:10.

50. Bradley RA: **A Biometrics Invited Paper. Science, Statistics, and Paired Comparisons**. *Biometrics* 1976, **32**(2):213-239.
51. Arnesen T, Nord E: **The value of DALY life: problems with ethics and validity of disability adjusted life years**. *Lepr Rev* 2000, **71**(2):123-127.
52. Hosseinpoor AR, Bergen N, Kostanjsek N, Kowal P, Officer A, Chatterji S: **Socio-demographic patterns of disability among older adult populations of low-income and middle-income countries: results from World Health Survey**. *International Journal of Public Health* 2016, **61**(3):337-345.
53. 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**(1):198.
54. Ayuso-Mateos JL, Nuevo R, Verdes E, Naidoo N, Chatterji S: **From depressive symptoms to depressive disorders: the relevance of thresholds**. *British Journal of Psychiatry* 2010, **196**(5):365-371.
55. 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**. *The Lancet* 2007, **370**(9590):851-858.
56. Araya R, Zitko P, Markkula N: **The Impact of Universal Health Care Programmes on Improving 'Realized Access' to Care for Depression in Chile**. *Adm Policy Ment Health* 2018, **45**(5):790-799.
57. Chile MdSd: **Encuesta Nacional de Saud ENS Chile 2009-2010**, Ministerio de Salud de Chile edn; 2010.
58. C. KR, Gavin A, Daniel M, Bedirhan U, Hans-Ulrich W: **The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF)**. *International Journal of Methods in Psychiatric Research* 1998, **7**(4):171-185.
59. Markkula N, Zitko P, Peña S, Margozzini P, Retamal C. P: **Prevalence, trends, correlates and treatment of depression in Chile in 2003 to 2010**. *Social Psychiatry and Psychiatric Epidemiology* 2017, **52**(4):399-409.
60. Barendse MT, Oort FJ, Timmerman ME: **Using Exploratory Factor Analysis to Determine the Dimensionality of Discrete Responses**. *Structural Equation Modeling: A Multidisciplinary Journal* 2015, **22**(1):87-101.
61. Everitt BS: **An R and S-PLUS® companion to multivariate analysis.** London: Springer-Verlag; 2005.
62. Revelle W: **An introduction to psychometric theory with applications in R**. In. Edited by Project TP. Illinois, USA: The Personality Project; 2020.
63. Mulaik S: **Foundations of Factor Analysis**. New York: Chapman and Hall/CRC; 2009.
64. Humphreys LG, Montanelli Jr RG: **An Investigation of the Parallel Analysis Criterion for Determining the Number of Common Factors**. *Multivariate Behavioral Research* 1975, **10**(2):193-205.
65. Muthén B, Kaplan D: **A comparison of some methodologies for the factor analysis of non-normal Likert variables**. *British Journal of Mathematical and Statistical Psychology* 1985, **38**(2):171-189.
66. Bollen KA: **Structural Equations with Latent Variables**. United Satates of America: John Wiley & Son; 1989.

67. Hoyle RH: **Handbook of structural equation modeling**: Guilford Publications; 2012.
68. Satorra A, Bentler PM: **Ensuring Positiveness of the Scaled Difference Chi-square Test Statistic**. *Psychometrika* 2010, **75**(2):243-248.
69. Maerlender A, Flashman L, Kessler A, Kumbhani S, Greenwald R, Tosteson T, McAllister T: **Discriminant construct validity of ImPACT™: a companion study**. *Clin Neuropsychol* 2013, **27**(2):290-299.
70. Cronbach LJ: **Coefficient alpha and the internal structure of tests**. *Psychometrika* 1951, **16**(3):297-334.
71. Nunnally J, Bernstein I: **Psychometric Theory (3rd edition)**, 3rd edn. New York: McGraw-Hill; 1994.
72. Tavakol M, Dennick R: **Making sense of Cronbach's alpha**. *Int J Med Educ* 2011, **2**:53-55.
73. Rosseel Y: **lavaan: An R Package for Structural Equation Modeling**. *2012* 2012, **48**(2):36.
74. Salomon JA, Murray CJL, Üstün TB, Chatterji S: **Health State Valuations in Summary. Measures of Population Health**. In: *Health Systems Performance Assessment Debates, Methods and Empiricism.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 409-436.
75. Organization WH: **Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0)**; 2010.
76. Federici S, Bracalenti M, Meloni F, Luciano JV: **World Health Organization disability assessment schedule 2.0: An international systematic review**. *Disability and Rehabilitation* 2017, **39**(23):2347-2380.
77. Brooks R: **EuroQol: the current state of play**. *Health Policy* 1996, **37**(1):53-72.
78. Devlin NJ, Brooks R: **EQ-5D and the EuroQol Group: Past, Present and Future**. *Appl Health Econ Health Policy* 2017, **15**(2):127-137.
79. Nord E: **Uncertainties about disability weights for the Global Burden of Disease study**. *The Lancet Global health* 2015, **3**(11):e661-662.
80. Palazzo C, Yokota RTC, Ferguson J, Tafforeau J, Ravaud JF, Van Oyen H, Nusselder WJ: **Methods to assess the contribution of diseases to disability using cross-sectional studies: comparison of different versions of the attributable fraction and the attribution method**. *Int J Epidemiol* 2018.
81. Kolovos S, Bosmans JE, van Dongen JM, van Esveld B, Magai D, van Straten A, van der Feltz-Cornelis C, van Steenbergen-Weijenburg KM, Huijbregts KM, van Marwijk H *et al*: **Utility scores for different health states related to depression: individual participant data analysis**. *Qual Life Res* 2017, **26**(7):1649-1658.
82. Zarate V, Kind P, Valenzuela P, Vignau A, Olivares-Tirado P, Munoz A: **Social valuation of EQ-5D health states: the Chilean case**. *Value Health* 2011, **14**(8):1135-1141.
83. Murray CJL, Ozlatin E, Tandon A, Solomon JA, Sadana R, Chatterji S: **Empirical Evaluation of the Anchoring Vignette Approach in Health Surveys**. In: *Health Systems Performance Assessment Debates, Methods and Empiricism.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 369 - 400.

84. Masyn KE: **Measurement Invariance and Differential Item Functioning in Latent Class Analysis With Stepwise Multiple Indicator Multiple Cause Modeling**. *Structural Equation Modeling: A Multidisciplinary Journal* 2017, **24**(2):180-197.

Chapter 3: A practical approach to calculate the Burden of Disability due to Diseases using data from National Health Surveys

3.1. Introduction

In the first chapter, we observed that most of the studies which used effective coverage to measure universal health coverage or to assess the performance of health systems, nonetheless fail to follow the formulation made by Shengelia et al. [1]. Hardly any of the previous studies that analysed effective coverage used direct measurements of the health gains associated with healthcare interventions.

In the second chapter, I discussed the concepts of health, disability, health states, and health-disability-utility weights, which are essential elements that must be considered to measure health gains. Also, I argued for the convenience of the conceptual framework of the International Classification of Functioning, Disability and Health (ICF-2) from the World Health Organisation (WHO). In addition, I showed how, by a latent variable approach, health-disability weights for health conditions could be calculated in one step, without identifying specific health states. In chapter two, I also assessed the psychometric properties and dimensionality of a health state description questionnaire included in a national health survey carried out in Chile, which I am using to test the proposed methodology to estimate effective coverage.

In this chapter, I present a new approach for calculating the fraction of the burden of disability attributable to diseases. This approach is especially developed to be implemented for continuous outcomes – for a latent variable of disability, for example. This fraction is a crucial component of the method I developed to calculate effective coverage, as it will be based on the burden of disability estimated under different hypothetical scenarios for the coverage of healthcare interventions.

Because the goal is to estimate the burden of disability, in the second section of this chapter I will briefly introduce the concept of the burden of disease and how it has been traditionally calculated. This background will allow a greater appreciation of the new approach I am proposing. In the same section, I will introduce the concept of attributable fractions, the different approaches available to estimate it, and the context in which this estimator is commonly used. This background will also be required to adequately understand all the indicators derived from the proposal to estimate effective coverage found in chapters four and five.

In the third section, I will present the methodological proposal for calculating attributional fractions onto a continuous outcome, and two complementary approaches to estimate the burden of disability. I will also show how these procedures can be applied to different domains of disability.

In the fourth section, I will present the results from the methodological proposal applied to a selected group of chronic non-communicable diseases using data from the Chilean National Health Survey 2009-2010. In addition, I will explore how sensitive the results are to different initial specifications.

In the fifth section, I will discuss the methods and results by comparing them with currently existing approaches.

Finally, in the sixth section, I will provide conclusions and a summary of this chapter's contents.

3.2. Background

The two key metrics I will discuss in this chapter are: (1) the burden of disease and (2) the attributable fraction. I propose to gather both metrics in only one expression, which will account for the burden of disability attributable to diseases. To do that, firstly, I will provide a relevant contextual and conceptual background about the burden of disease.

3.2.1 Burden of disease

Generically, the term burden of disease is used to describe the impact of a health condition on a particular outcome, such as mortality, disability or costs. However, in the last twenty years the term has been used, almost unequivocally, as a synonym for Disability Adjusted Life Years (DALY), a metric that accounts for the number of years lost to disability and mortality as a consequence of diseases.

The metric was firstly launched by the World Bank (WB) in its annual World Development Report for 1993, which addressed the topic of investment in health. In that report, for the first time, estimates of DALYs for the year 1990 in all regions of the world were reported, including 107 diseases [2]. The main developers of the metric were J. Christopher Murray and Alan Lopez [3], who later published the results of their study under the name Global Burden of Disease Study (GBD) [4]. Updates of the GBD were launched by the WB jointly with the WHO in 2006, using data from 2000-2002 [5], and by the WHO alone in 2008, using data from 2004 [6]. Between 1993 and 2008, several countries carried out their own national burden of disease studies [7-12]. The DALYs were also used as one of the indicators to measure the performance of national health systems in the World Health Report of 2000 [13].

After 2008, the GBD studies were undertaken by the Institute of Health Metrics and Evaluation (IHME), hosted at the University of Washington, in the US. There have been reports in 2012 (using data from 2010) [14], in 2015 (data from 2013) [15], in 2016 (data from 2015) [16], and in 2018 (data from 2017) [17].

The metric of DALYs can be expressed as:

$$DALY = YLD + YLL$$

where YLD is the number of years lost to disability, and YLL is the standard expected years lost to premature mortality. The YLL can be calculated through the equation:

$$YLL = \sum_{x=0}^{x=L} d_x(L - x)$$

where *d* is the number of deaths at age *x* caused by the disease, and *L* is the potential limit to life [18]. Essentially, the YLLs are the total number of years lost as a consequence of a disease, calculated using the age of death and a standard value of life expectancy.

On the other hand, the YLD at an individual level can be expressed as:

$$YLD = \sum_{x=0}^{x=M} I_x DW$$

where I is the number of incident cases of a disease at the age x, D is the average duration of the disability, W is the disability weight (a concept introduced in chapter two), and M is the maximum age of the population [3]. Since the prevalence of a disease can be calculated by multiplying the incidence by its duration, this equation can be simplified, as it was in the latest reports of the IHME:

$$YLD = \sum_{x=0}^{x=M} P_x W$$

, where *P* is the prevalence of the disease at age *x* [19].

In the history of GBD studies, several methodological advances have been produced, especially after the leadership of the IHME during the last ten years. However, implementing these advances is complex and challenging when a single country wants to calculate their burden of disease. The challenges and complexities are listed below.

3.2.1.1 Complexity 1: Parameters are based on predictions using foreign data

The current strategy of the IHME to calculate the YLD globally is essentially based on a meta-regression model used to predict the prevalence of diseases at each stratum of age, sex, year, and country (or subnational areas). Depending on the disease, the model can include variables such as gross domestic product, a composite index about the performance of health systems, or a composite index of socio-demographic development in the country, among other variables specific for each health condition. As a result, the predicted prevalence of a particular disease for a certain country and year is not only based on information about the specific country – for example, national health surveys – but also on information from other countries and predictors of prevalence. This means that, if a single country is interested in replicating the IHME's estimates, information from all the predictors for different countries included in the model must be collected. Therefore, the IHME's approach is not particularly suitable for countries that want to calculate their own burden of disease. Still, it has the advantage of predicting the prevalence of a disease, even in countries where they have never been studied before.

3.2.1.2 Complexity 2: Knowing the relative weight of each disease first requires studying hundreds of other diseases

The first study of GBD included 107 diseases, but the latest studies have expanded to 359 diseases. This improvement is not without challenges for individual countries, especially when there is a need to estimate the relative weight of each disease in the total burden. For example, for cases of major depressive disorder in Chile, the IHME estimated 113,136 YLD [79,326 - 152,997] in 2010, which is equivalent to 5.8% [4.7 - 7.0] of the total burden of disability for the country. In order to calculate this fraction, it was necessary to estimate the total number of YLD produced by the whole list of 359 diseases in Chile, which according to the IHME was 1,952,271 YLD [1,466,137 - 2,527,004] [extracted from https://vizhub.healthdata.org/ gbd-compare/ in September 2020] to reach the 5.8% fraction (113,136/ 1,952,271).

3.2.1.3 Complexity 3: Source and procedures behind disability weights

In chapter two, I mentioned that the IHME used a paired comparison procedure to value different health states and build the disability weights. However, each disability weight for a

particular disease is the result of a combination of specific disability weights from different sequels – called severities in early versions of GBD studies – associated with each health condition. For example, in the case of a major depressive disorder, the IHME describes three sequels (mild, moderate or severe episodes), each one with a different disability weight (0.145, 0.396, 0.658, respectively) [20]. Using multiple sequels for describing a disease reflects the interest in capturing adequately the different health states involved in each disease. However, the current list of sequels used by the IHME surpasses 3,400 [21].

The data on disability weights used by the IHME in the latest versions of the GBD's studies combined information collected in four countries (Bangladesh, Indonesia, Peru, and Tanzania) through household surveys, as well as one nationally representative telephone survey carried out in the US between 2009 and 2010, where 108 health states were valued. The data was complemented with a web-based survey conducted between 2010 and 2011, involving more than 30,000 people from 167 countries. In this study, 220 health states were valued [22]. The results were also pooled together with a second web survey conducted in four European countries (Hungary, Italy, the Netherlands, and Sweden) in 2013, where 255 health states were explored [20, 23]. The disability weights for all the sequels are based on the valuation of these health states.

Despite the wide representation – with a high number of countries participating in the estimation of disability weights – these studies used different selection criteria for participants in each sample, as well as different methods for collecting the data (i.e., face-to-face, telephone, and website surveys). Concerns have also been raised about the appropriateness of these disability weights for each different country, especially those not involved in the studied samples [24-27]. Other authors have criticised a lack of transparency in the methods used by the IHME to create the disability weights [28].

3.2.1.4 Complexity 4: Adjustment of disability weights by comorbidities

The valuation processes of health states by paired comparison allows comparing only two states each time, which preclude a direct adjustment by comorbidity. Therefore, it must be performed indirectly.

The procedure currently employed consists of simulating a population of 20,000 individuals with a probability for each sequel equal to its estimated prevalence, assuming sequels are independently distributed. Then, the total disability weight for each simulant is calculated by applying the respective disability for each sequel: one minus the product of one minus each disability weight. The final disability weight for each sequel corresponds to the average of all the simulants. However, the prevalence of the sequels is uncertain, and to account for this the simulation process is repeated one thousand times, according to distributional assumptions [19, 29]. This procedure is computationally intensive since it generates twenty million simulations for each sequel, considering just one stratum of age, sex, country and year. These estimations can only be done using special computers (for example, in Chile, there is only one computer able to perform such calculations within a reasonable time). Alongside these technical difficulties, the assumed independence between probabilities for each sequel can also be questioned, although tests on real data have shown good reliability [29].

3.2.1.5 Other complexities

There are several other complexities included in the current methodology developed by the IHME that are beyond the scope of this section of the chapter. For example, there are procedures to ensure the consistency of prevalence data, and other parameters such as disease duration, incidence, or mortality across different ages in the same population. Furthermore, several diseases require variations in the methodology to estimate YLD (e.g., HIV/AIDS, injuries), including additional inputs to estimate the YLD for each sequel. Other elements, such as using a discount rate and a function giving different weights to YLD produced at different ages, have been sources of ethical concerns and have been removed from the latest versions of GBD studies [19].

In conclusion, there is no doubt about the massive improvements in the current models used to estimate the GBD, nor any doubt about their contribution to the understanding of global health. The IHME currently produces estimates for the burden of disease from 1990 onward, covering more than 250 countries and subnational territories, including information stratified by age and sex for 359 diseases in total. However, its methodological sophistication has made calculating DALYs, particularly the YLDs, more complicated. This, in turn, has moved these metrics beyond the capacities of national health agencies and academics. Some authors have suggested that the overwhelming amount of information generated at each

GBD study, and the frequency with which they are updated, is a source of discouragement for the strengthening of national health information systems, especially in middle- and low-income countries. This has led to a situation where the generation and processing of essential health information have been transferred from lower-income to higher-income countries. It has been argued that the leadership of the IHME is not only a technical matter but a political and normative matter too [30].

3.2.2 The attributable fraction

The second metric involved in the procedure I will present in this chapter, as well as in chapters four and five, is the attributable fraction (AF). This is a generic concept which encompasses a set of indicators combining absolute and relative measures of effect.

3.2.2.1 The beginning of the attributable fraction

The AF has been documented in the academic literature since 1951, when R. Doll calculated the number of lung cancer cases that could have been averted in London if the population's tobacco consumption was null, using data from a case-control study [31]. Similarly, in the US in 1951, J. Cornfield published an analogue approach to estimate lung, breast and cervix cancer deaths due to different risk factors (i.e., tobacco consumption, number of children, and age), using information from several case-control studies [32]. Initially this approach was called the 'Cornfield method' and was routinely used by the National Cancer Institute in the US. Later, in 1953, M.L. Levin formally derived the expression that today we call the population attributable fraction (PAF), in terms of rate ratios rather than rate differences, as was originally proposed by R. Doll and J. Cornfield. Essentially, M.L. Levin proposed (according to A. Leviton [33], and C. Poole 2015 [31]) that the total incidence rate (I_T) of cases (e.g. lung cancer deaths) can be decomposed following:

$$I_T = p(I_1) + (1 - p)I_0$$

where p is the prevalence of the exposure to a certain risk factor (e.g. smoking), I_1 the incidence rate (e.g. lung cancer) among the exposed, and I_0 the incidence rate among the non-exposed. Therefore, the total incidence rate is described as the sum of the incidence

rates in exposed and non-exposed, weighted by the prevalence of the exposition. This equation can also be expressed as:

$$I_T = p(I_1 - I_0) + I_0$$
$$I_T = p(ID) + I_0$$

where *ID* is the incidence rate difference, and

$$PID = I_T - I_0 = p(ID)$$

where PID is the population incidence rate difference, which corresponds to the number of incident cases attributable to the exposition. Then, the PAF (originally called the 'proportion attributable' by M.L. Levin) was defined as:

$$PAF = (I_T - I_0) / I_T \qquad \text{(Eq.3.0)}$$
$$PAF = PID / I_T$$
$$PAF = p(ID) / [\, p(ID) + I_0 \,]$$

which is equivalent to the fraction of incident cases attributable to the exposure. The numerator and denominator of this equation can be divided by I_0, resulting in:

$$PAF = \frac{p(IR - 1)}{p(IR - 1) + 1} \qquad \text{(Eq.3.1)}$$

where IR is the incidence rate ratio. Additionally, M.L. Levin calculated the proportion of cases among the exposed accounted for by the rate difference, which would later be called the exposed attributable fraction (EAF):

$$EAF = ID / I_1$$

The numerator and denominator of the EAF can also be divided by I_0, resulting in IR – 1, or by I_1, resulting in:

$$EAF = \frac{IR - 1}{IR} = 1 - \frac{1}{IR} \qquad \text{(Eq3.2)}$$

which is understood as the proportion of excess cases among the exposed.

3.2.2.2 Early development of AF: A conflicting terminology

The history of PAF and related measures of the AF is confusing in terms of names and concepts. For example, in 1960 the concept of PAF and EAF were presented for the first time in an epidemiological textbook [34], without providing any specific names, whilst Levin's concepts of ID and PID were called 'attributable risk' and 'attributable community risk' respectively. Moreover, seemingly under the influence of J. Cornfield, that textbook was oriented towards specific measures to calculate attributable cases of cancer, using information from case-control studies [31].

In 1970, in a new textbook of epidemiology from the same authors, introduced the concept of 'attributable community risk' and 'population attributable risk'. Then, in 1971, P. Cole and B. MacMahon published on how to calculate the PAF and AF using 'relative risk' derived from case-control studies on cancer, in contexts where the rate of the disease was unknown in the target population. They derived the same equations as M.L. Levin, without apparently being aware of that previous work. In their publication, P. Cole and B. MacMahon referred to the PAF and EAF as 'population attributable risk per cent' and 'attributable risk per cent', respectively [35].

More confusion in terminology was raised during the 1970s when A.M. Lilienfeld, an influential epidemiologist, suggested that M.L. Levin named the PAF 'attributable risk'. At the same time, he preferred to use the term 'population attributable risk' when referring to the PAF [36], even though MacMahon and Pugh had used that term to denominate the PID [31]. Leviton also referred to the PAF as 'attributable risk' [33]. In addition, in 1974, the PAF was described as an 'etiologic fraction' by O. Miettinen [37], a concept that later became more clearly differentiated from AF measures. Miettinen also explored the relationship between PAF and EAF, providing an important development for this metric (PAF = pcEAF; where pc is the proportion of cases in the higher level of exposure).

3.2.2.3 The Population Attributable Fraction (PAF)

Some relevant progress in the AF metric occurred in 1976 and 1979 when the concept of PAF was applied to categorical levels of exposures. The modified PAF equation, according to S.D. Walter and R.P. Ouellet et al. is [38, 39]:

$$PAF = \frac{\sum_{i}^{n} p_i\,(RR_i - 1)}{\sum_{i}^{n} p_i\,(RR_i - 1) + 1} \qquad \text{(Eq.3.3)}$$

where *i* is the level of exposure, and n is the total number of levels of exposure for the risk factor. This equation keeps the structure of equation 3.1, but the numerator represents the sum of cases in excess from each level, divided by the total number of cases in the entire population. It is worth noting that S.D. Walter called the PAF 'attributable risk', while R.P. Ouellet et al. used the term 'population attributable risk'.

In 1976, S.D. Walter, also proposed a method to estimate the variance for PAF, and how to adjust PAF by confounding factors, as well as how to include the interactions between risk factors in the calculations, all applied to data arranged in a contingency table [38].

3.2.2.4 The Population Impact Fraction (PIF)

In 1982, Morgenstern et al. extended the use of the PAF to situations in which the absence of exposition was not the basal level of comparison, an underlying assumption in equations 3.0, 3.1, 3.2, and 3.3. The authors used the term potential impact factor for this new metric [40]. Their starting point was the Eq3.0 (i.e. PAF = $(I_T - I_0)/I_T$), defining I_T for *n* categorical level of expositions of the risk factor:

$$I_T = \sum_{i}^{n} p_i I_i$$

where *i* is the strata of the exposition. As a consequence, the PAF (Morgenstern used the term 'AF') from Eq3.0, can be defined directly in the following way:

$$PAF = \frac{\sum_{i}^{n} p_i I_i - I_0}{\sum_{i}^{n} p_i I_i}$$

After dividing the numerator and denominator by I_0, the resultant equation is:

$$PAF = \frac{\sum_{i}^{n} p_i RR_i - 1}{\sum_{i}^{n} p_i RR_i} \quad \text{(Eq.3.4)}$$

which is equivalent to Eq.3.3. This expression can be accommodated to cases where the causal contrast in the numerator of the PAF is no longer I_T against I_0 (remember that I_0 is the category where the exposition is absent), but I_T against the resultant incidence rate from a counterfactual distribution of the risk factor. In this way the PIF can be defined as:

$$PIF = \frac{\sum_{i}^{n} p_i RR_i - p_i' RR_i}{\sum_{i}^{n} p_i RR_i} \quad \text{(Eq.3.5)}$$

where p' describes the counterfactual scenario of the distribution of the risk factor. As a consequence, the PAF corresponds to the maximum PIF that an intervention addressed to remove a risk factor can expect [40]. The distinction between PAF and PIF is relevant, since both concepts can be used to measure effective coverage (see chapter four).

The PAF can be described through several formulations. However, for the sake of clarity, equation 3.0 reflects the most basic concept of PAF, while equation 3.1 is its best-known formulation. Equation 3.3 and 3.4 are equivalent and represent the PAF extended to risk factors with a categorical level of expositions. Finally, equation 3.5 is a more general expression than the PAF and considers different counterfactual levels of the risk factor as a comparator. Continuous expositions can also be resolved using equations 3.4 and 3.5,

although they require integrating the distribution of the exposition on a continuous function of the relative risk [41].

3.2.2.5 Latest progress in AF

One of the latest methodological approaches to the calculation of the PAF was developed by P. Bruzzi et al. in 1985. The authors proposed the use of regression logistic models with data from case-control studies for calculating the PAF adjusting by multiple risk factors and confounders [42]. This procedure proved to be more flexible and straightforward than previous approaches, based on contingency tables [38]. Of note, Bruzzi et al. used the term 'population attributable risk per cent' or 'etiologic fraction' interchangeably to refer to the concept I am calling PAF. In addition, in 1993, S. Greenland and K. Drescher proposed a maximum likelihood estimator for PAF from logistic models [43].

3.2.2.6 Agreeing on terminology

Because different terms have been employed to denominate similar concepts, I will use the terminology suggested by K. Rothman et al. in 2008. According to these authors, the measures of AF derive from the combination of an absolute causal effect measure (i.e., rate difference, risk difference) and a relative one (i.e., rate ratio, risk ratio, jointly denominated relative risks). This can be noticed in equation 3.2, where EAF = (IR – 1) / IR*, which is equivalent to $(I_1 - I_0)/ I_1$, and the numerator is an absolute expression of effect. However, on dividing by I_1 it is transformed into a relative one. When the measure of EAF is based on rates, K. Rothman et al. used the specific term of rate fraction, whilst if it was based on risk, they preferred the term risk fraction. Both measures of EAF are also named 'excess fractions', as opposed to the concept of etiological fraction. While the EAF informs the fraction of events (or cases) in excess among the exposed, the PAF informs that fraction for the entire population (exposed and unexposed to the risk factor) [44].

* *K. Rothman et al. call the exposed attributable fraction (EAF) just the attributable fraction (AF). For convenience, I have opted to use the term AF as a generic term that encompasses all the concepts related to effect measures that combine relative and absolute metrics in only one expression. That is, AF includes the measures of EAF and PAF.*

3.2.2.7 Relationship between AF and Etiological Fraction

The distinction between AF measures (or excess fractions) and the etiological faction is relevant. Since several risk factors (smoking, pollution, age, and so on) can combine to generate an event of interest (e.g. death by lung cancer), and also since the time-lapse from exposure to event can differ (i.e. the etiological time), the cases attributable to a certain exposition must be interpreted only as the lowest boundary of the total cases caused by such an exposition. [45]. This is because, when a certain exposition is removed, the number of events or cases averted can be partially replaced by new events or cases arising from other expositions (during the time window of the study). Consequently, in an epidemiological study, we can only observe the reduction in cases that are not replaced by cases from other causes. To calculate the etiologic fraction, it is necessary to include strong assumptions about the causal mechanisms that lead to an event of interest. Unfortunately, through epidemiological studies (or even randomised controlled trials), it is often not possible to find out the fraction of events or cases that can be replaced by the presence of other factors [44, 45]. Therefore, the AF must be interpreted as the lowest possible etiologic fraction. In other words, the etiological fraction can never be lower than the AF.

3.2.2.8 Causality mechanisms and AF: Metric consequences

Another element that emerges from the causal analysis of multiple risk factors and the incidence of cases is that different risk factors can share similar causal mechanisms or be located in different positions in the same causal chain. For example, when body mass index (BMI) is explored as a risk factor for cardiovascular deaths, the calculated PAF also includes the hypothetical reduction of other intermediate risk factors such as high blood pressure and hypercholesterolemia derived from a high BMI. This means that the PAF for BMI in part accounts for the PAF due to high blood pressure and hypercholesterolemia separately. Moreover, one risk factor can modify its effect on the outcome when coexisting with another risk factor, reflecting complex plausible causal interactions. As a consequence, the PAF from different risk factors for a single outcome cannot simply be added, and when that is done on a large list of risk factors, it is common to obtain a total PAF higher than 100% of the cases (i.e. the risk factors jointly account for more cases than actually exist) [46]. This consideration and its implications will be discussed later in the chapter.

3.2.2.9 Two areas where the AF metric has not been well developed

The metric of AF was closely related to the study of risk factors, especially in the context of case-control studies. This has several implications. Firstly, and most importantly, the outcome used to calculate the PAF is traditionally the number of incident cases or events (i.e. new cases of a disease or deaths caused by a disease). In other words, a discrete outcome. One of the few studies that used AF for a different outcome was carried out by Tanusepruto et al. and published in 2015. The authors used a Cox proportional hazard model to estimate deaths attributable to tobacco consumption via a survival analysis [47]. The lack of studies applying continuous outcomes is relevant, given that my proposal for the burden of disability attributable to diseases is based on the use of continuous values of disability.

Secondly, there are very few studies that have evaluated the impact of diseases on disability, rather than on morbidity or mortality, which is the traditional approach in AF studies. I am proposing to study the burden of disability attributable to diseases, treating disease as a risk factor and the disability as the outcome. Palazzo et al. [48, 49] are among the few authors that have explored the contribution of diseases to disability through AFs. However, following the tradition of risk factors, the authors defined disability as a dichotomous event (i.e. disabled or not disabled) and, by implementing the Bruzzi et al. approach, they calculated the number of disabled people in excess attributable to different diseases [48, 49]. Unfortunately, as discussed in chapter two, assuming disability as a categorical construct goes against current theoretical frameworks for health state valuations [50].

3.2.2.10 A final element of context in the metric of AF

To conclude this section, I would like to mention the largest studies that have been conducted using the AF metric. They are the global 'Comparative Risk Assessment' (CRA) studies, developed as part of the GBD project mentioned previously. The first global CRA, where the number of deaths, YLDs, YLLs and DALYs attributable to ten major risk factors were estimated for eight of the world's regions was published in 1997 [4]. This study was updated in 2002, exploring 26 risk factors for 14 of the world's regions [51-53], and later in 2006, where 19 risk factors were studied for 192 countries [5]. A fourth update was carried out by the WHO in 2008 [54]. Since 2012, updates of CRA studies are published simultaneously with each new GBD report [55-57]. The last CRA study, from 2018, included 84 risk factors applied

to 354 diseases, and it gives estimates for 195 countries in total, for all the years between 1990 and 2017, stratified by age and sex [58].

Basically, these reports used the following approach to estimate the burden of disease attributable to certain risk factors:

$$\text{DALY attributable} = (\text{YLD} * \text{PIF}_{\text{incident cases}}) + (\text{YLL} * \text{PIF}_{\text{mortality}})$$

In other words, this represents the sum of the YLDs and the YLLs attributable to the risk factor. Notice that the PIF for YLD and YLL are different. That happens not only because the incidence of new cases is not necessarily equal to the mortality rate, but also because the relative risk can vary. In the context of CRA, the use of the PIF is preferred to the PAF, because it is a more general expression that allows us to compare different counterfactual distributions of the risk factor, as mentioned above. This is convenient for some exposures, where a total absence cannot be expected (e.g., BMI, blood pressure, etcetera). It Is relevant for this thesis to point out that the CRA usually considers four types of counterfactuals: 1) the theoretical minimum (usually equivalent to zero exposure), 2) the plausible (i.e. a counterfactual achieved by a reference to experience), 3) the feasible (i.e. the counterfactual that may realistically be expected in the context of the study), and 4) a cost-effective counterfactual (i.e. a choice using a cost-effective criterion) [46]. I will discuss this topic further in chapter four.

3.2.3 Synthesis of the background

The burden of disease (i.e. DALYs) and AF are very well-known epidemiological metrics. The disability component of the burden of disease is the YLD, which summarises the experience of living with disability due to a disease. However, as was described above, to estimate the relative impact of one disease on the total burden of disability in a population is a challenging task, and usually beyond the resources of national ministries of health or traditional academic institutions. Moreover, the disability weights used in the YLD have been criticised for their potential lack of relevance to local contexts, and also by the assumptions involved in the procedure used to adjust them to multimorbidity.

On the other hand, AF has been extensively used in the context of the assessment of risk factors on the incidence of diseases and deaths. However, very few studies have explored its use in accounting for the burden of disability due to diseases. Moreover, these studies do not consider disability as a continuous attribute, which is against the theoretical framework of health state valuations.

Having reviewed the basis and limitations of the available metrics for the burden of diseases and AF, I can now introduce a new approach where I intend to combine both metrics in one. This new metric should resolve several of the complexities and limitations already presented, and also serve as the basis for calculating the effective coverage of healthcare interventions.

In the following section of this chapter, I will present the proposal to measure the burden of disability attributable to diseases, in the context of cross-sectional data. This proposal has the advantage of generating disability weights adjusted by comorbidity, suitable to the local context, and also calculating the relative burden of disability through a straightforward approach. The procedure also has the merit of using the AF on a different outcome to those traditionally used in CRA studies.

3.3. Methods

3.3.1 General framework for measuring the burden of disability attributable to diseases

I propose to estimate the burden of disability due to disease, in the context of cross-sectional data, setting two scenarios: 1) the actual scenario which corresponds to the observed disability in the whole population (D); and 2) the counterfactual scenario (D') in which the disability attributable to a particular disease is not present. Therefore, $D' = D - D_A$, where D_A corresponds to the attributable disability, or, in other words, the excess of disability attributable to the disease. The burden is estimated according to the following relationship:

$$\text{Burden of disability due to a disease} = 1 - \frac{D'}{D} = \frac{D_A}{D} \qquad \text{(Eq.3.6)}$$

Notice that D_A/D is an expression very similar to the equation proposed by M.L. Levin in 1953 ($[I_T - I_0] / I_T$) to calculate the PAF (see equation 3.0). However, in this case, the total incidence (I_T) is represented by the actual scenario of disability in the entire population, and I_0 describes the counterfactual scenario, which depicts the absence of the disease.

There are, at least, two different approaches that can be followed to calculate the burden of disability expressed in equation 3.6. One is based on the population average-level and one on individual-level estimates, which I describe in detail below.

3.3.2. Population average-level approach

The actual amount of disability in the population (D) can be decomposed into the sum of the disability of people without the disease (D_0) and the disability of people with the disease (D_1). Following a *population average-level* approach, D_0 can be specified as $E(D_0)(1 - P)*N$, where $E(D_0)$ is the expected value of D_0, P is the prevalence of the disease and N the number of individuals in that population. Similarly, D_1 can be expressed as $E(D_1)PN$, and D_A as $E(D_A)PN$, leading to the following equation, where N is cancelled:

$$\text{Burden of disease on disability} = \frac{\mathrm{E}(\mathrm{D_A})\mathrm{P}}{\mathrm{E}(\mathrm{D_0})(1-\mathrm{P}) + \mathrm{E}(\mathrm{D_1})\mathrm{P}} \qquad \text{(Eq.3.7)}$$

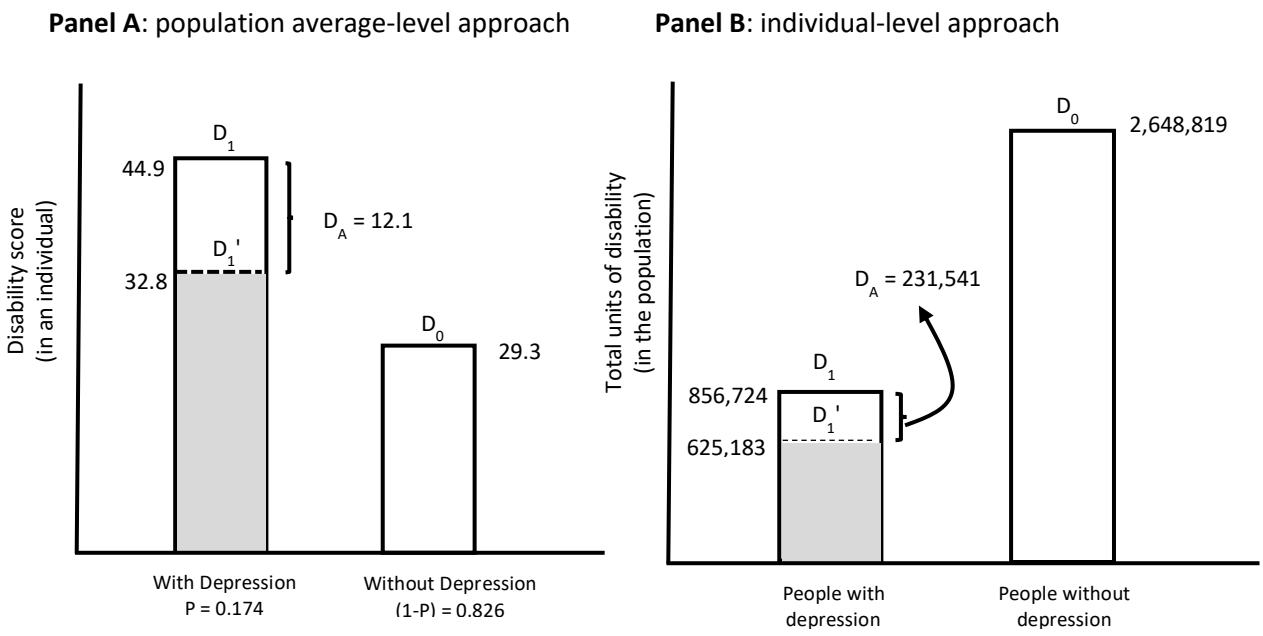
The expected values of D_0 and D_1 can be estimated through the mean of the score of disability in individuals without and with the disease, respectively. Also, they can be estimated by predictions from a linear regression model, assuming normality in the distribution of residuals, applied to the data of each respective population and setting the values of several covariates at the mean value:

$$E(D_0) \approx \widetilde{D}_0 = \hat{B}\,\bar{X}_0 \;; \qquad \text{and} \quad E(D_1) \approx \widetilde{D}_1 = \hat{B}\,\bar{X}_1 \qquad \text{(Eq.3.8)}$$

where $\widetilde{D}_0$ and $\widetilde{D}_1$ are predicted values of disability, $\hat{B}$ is the vector of estimated regression coefficients, and $\bar{X}$ is the transposed vector containing the mean population value of covariables from individuals without ($_0$) or with ($_1$) the disease.

The disability attributable (D_A) can be estimated using similar linear regression models, as mentioned above, applied to the whole population. In a regression model, D_A corresponds to the regression coefficient from the variable that marks the presence of the disease (see chapter two). However, in a more complex model, which could include interaction terms, it is easier to estimate D_A as $\widetilde{D}_1 - \widetilde{D}_1$', where $\widetilde{D}_1$' is the predicted value of disability using the regression coefficients from the model for people with the disease, assuming they do not have the disease (i.e. replacing with 0 values to the variable that mark the presence of the disease), keeping the means of covariates constant. This is a similar procedure, to that developed by P. Bruzzi et al., in 1985, applied to case-control studies using logistic regression models (although they were interested in modelling the probability to be a case for different strata of risk exposition) [42]. A conceptual representation of the parameters is shown in Figure 3.1., Panel A.

Figure 3.1. Conceptual representation of parameters involved in the calculation of the burden of disease on disability, using a population average-level approach (Panel A), and a individual-level approach (Panel B).



Data used in this example is extracted from Table S3.1 (P: prevalence), and from Tables 3.1 and 3.3 (D_0, D_1, D_1', and D_A: disability of people without the disease, disability of people with the disease, disability of people with the disease assuming they have not got it, and attributable disability, respectively), according to depression, model 1, population average-level approach. Chilean National Health Survey 2009-2019 (n=4.447).

The use of regression models allows us to adjust the estimation of the D_A – and subsequently of the burden – for confounders such as comorbidities, as well as provide enough flexibility to test interaction terms between covariates.

3.3.3 Individual-level approach

Additionally, the procedure to estimate the burden of disability due to diseases, in the context of cross-sectional studies can be implemented through a second approach using predicted values at the *individual-level*. The procedure is equivalent to equation 3.7, but using individual predictions from a linear regression model.

$$\text{Burden of disease on disability} = \frac{\sum_{k=1}^{m} (\widetilde{D}_k - \widetilde{D}_k{}') W_k}{\sum_{i=1}^{n} \widetilde{D}_i W_i + \sum_{k=1}^{m} \widetilde{D}_k W_k} \quad \text{(Eq.3.9)}$$

where *n* is the total number of individuals *i* without the disease; *m* is the total individuals *k* with the disease; $\widetilde{D}$ is the predicted disability with the use of a regression model but applied to the dataset of individuals with ($_k$) or without ($_i$) the disease; $\widetilde{D}'$ is the predicted disability in people with the disease assuming they do not have it; and *W* is the probability weights of a population survey (i.e. 1/ probability to being chosen). The numerator of the equation corresponds to D_A, whilst the denominator depicts the total disability in the entire population (see Figure 3.1, Panel B).

3.3.4. Analysis by sub-domain of disability

In chapter two, I discussed the fact that disability and health status could be represented through different domains of functioning. According to that principle, both approaches to calculating the burden of disability may be applied to sub-domains of disability. This can be implemented by replacing the outcome of disability with the score representing the sub-domain of interest. A complementary approach would calculate the fraction of the whole DA due to a disease that is attributable to a sub-domain of disability. More detail is provided in the supplementary material (see 'Additional methods S3.1').

3.3.5. Data used to test the procedures

To test these proposals to calculate the burden of disability attributable to diseases, I used data from the second Chilean National Health Survey (Ch-NHS), the same dataset that was introduced in Chapter Two (see section 2.3.2). In that chapter, I also presented the questionnaire used to measure disability and its psychometric properties. In summary, the disability score ranges between 0 and 100, where 0 represents the total absence of difficulty

in carrying out any of 18 activities or participations (aggregated into nine sub-domains of disability), and 100 reflects total impossibility or extreme difficulty in carrying them out.

3.3.6 Diseases and covariables

To assess the burden of disability attributable to diseases, I selected five non-communicable conditions available in the Ch-NHS, all of which are highly prevalent, a common cause of mobility issues addressed by health services, and cover both physical and mental health disorders. They are hypertension, diabetes, depressive episodes, chronic respiratory symptoms and chronic musculoskeletal pain. The identification of cases with the first three conditions in the Ch-NHS was described in chapter two, section 2.3.3. Chronic respiratory symptoms and chronic musculoskeletal pain were added to the present analysis to enrich the comparison of results between diseases and also to demonstrate the effect of adjustment by comorbidity better.

Chronic respiratory symptoms were identified using items from a previous international study that aimed to determine the prevalence of respiratory disorders in Latin American countries [59]. Chronic respiratory symptoms were defined as the presence of a cough or phlegm unrelated to a cold, almost every day for at least three months of the year, during the last two years or longer. Chronic musculoskeletal pain was measured using the Community Oriented Programme for the Control of Rheumatic Disease Core Questionnaire, which was translated and locally validated [60]. Chronic musculoskeletal pain was defined as 'pain, stiffness, sensitivity or bone, muscle or joint swelling' located in one of 22 different body regions during the last seven days and lasting for more than three months, without a traumatic cause.

Additionally, for purposes of adjustment of regression models, age, sex and educational level variables were also used in the analysis. Education background was categorised in three levels: <8 years, 9-12 years, and >12 years of formal education, corresponding to primary, secondary, and high school or higher respectively in Chile.

3.3.7. Statistical analyses

I compared results using three linear regression models for the population average- and individual-level approaches: i) unadjusted (model 0); ii) adjusted for sex, age, education level and comorbidities (model 1); and iii) further adjusted to include two-way interaction terms for sex and comorbidities, and age and comorbidities (model 2). A significant interactions term allowed the attributable disability of a disease (i.e. D_A) to differ by sex and age and is therefore useful for exploring the extent that they may affect the estimate of the burden. The inclusion of interaction terms also allowed me to assess the approaches when applied to more complex regression models. The final two-way interaction parameters were derived from a backward stepwise selection procedure [61].

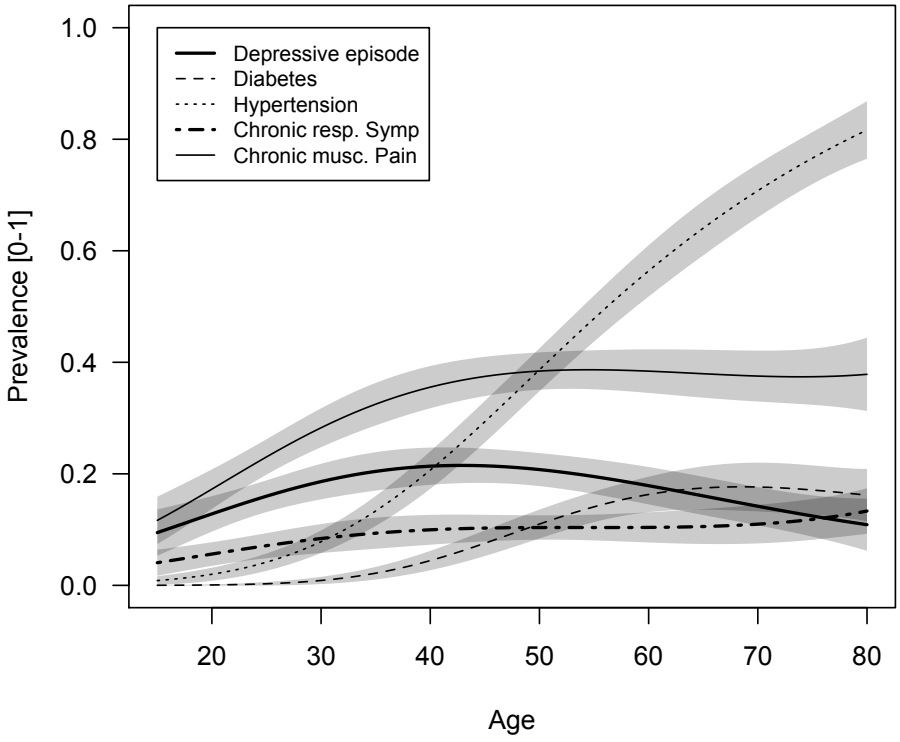
Analyses of the burden of disability due to diseases were also performed, stratified by sex and omitting this variable in the regression models presented above. Additionally, the burden was calculated by every year of age. Prevalence of disease according to age was first modelled using logistic regression models, adding quadratic and cubic terms for age when necessary. In the case of age, the stratified burden is reported without and with adjustment by the weight of each stratum on the total population disability. Direct estimates from the Ch-NHS are reported with 95% confidence intervals. For the attributable disability (DA) and the burden (equations 3.7 and 3.9), the uncertainty was estimated by simulating 10,000 values, assuming Normal distribution for parameters from regression models and Beta distribution for prevalence. In addition, quantile 50% is informed, whilst quantiles 2.5% and 97.5% are reported as uncertainty intervals [62].

The analysis was conducted using the statistical software R 3.5.1 and its package survey. The main functions for R created for this study are available in the supplementary material ('Main functions for R S3.1'). The database of the Ch-NHS can be downloaded from: http://epi.minsal.cl/condiciones-de-uso/ (consulted in August 2019).

3.4. Results

Descriptive statistics of the sample with the marginal prevalence of the selected health conditions are presented in the supplementary material, Table S3.1. In Figure 3.2, the estimated prevalence for the five selected health conditions are shown according to age.

Figure 3.2. Prevalence by age for five selected health conditions. Chilean National Health Survey 2009-2019 (n=4,447).



Chronic resp. Symp: chronic respiratory symptoms/ Chronic musc. Pain: chronic musculoskeletal pain

In Table 3.1 the regression coefficient is presented with 95% confidence intervals for disability from the unadjusted model (model 0); the adjusted model, which includes all health conditions and socio-demographic covariates (model 1); and the further adjusted model, including two-way interaction terms (model 2).

Table 3.1. Regression coefficients* for disability score [0 – 100] according to different diseases and covariates. Chilean National Health Survey 2009-2010 (n=4.447).

	Coefficients for Disability [0 - 100]					
	Model 0 (bivariate)		Model 1**		Model 2***	
	Coeff	**95% CI**	**Coeff**	**95% CI**	**Coeff**	**95% CI**
Intercept	-	-	16.0	[14 to 17.9]	10.8	[3.4 to 18.2]
Sex (women)	7.4	[6.0 to 8.8]	4.0	[2.8 to 5.2]	4.3	[3.0 to 5.6]
Age (x 10 years)	2.9	[2.5 to 3.2]	1.8	[1.4 to 2.1]	6.3	[0.9 to 11.7]
Education (> 12 years)	0.0	-	0.0	-	0.0	-
Education (9 - 12 years)	2.9	[1.2 to 4.5]	1.4	[-0.1 to 2.8]	1.5	[0.0 to 2.9]
Education (≤8 years)	10.7	[8.5 to 12.9]	4.5	[2.5 to 6.6]	4.5	[2.4 to 6.7]
Depressive episode	15.6	[13.8 to 17.4]	12.1	[10.5 to 13.8]	7.5	[3.1 to 11.9]
Diabetes	11.0	[8.5 to 13.5]	4.4	[2.1 to 6.7]	5.9	[1.7 to 10.0]
Hypertension	7.6	[5.9 to 9.3]	1.5	[-0.1 to 3.1]	4.1	[-2.0 to 10.1]
Chronic respiratory symptoms	8.5	[5.8 to 11.2]	6.1	[3.7 to 8.5]	10.8	[5.0 to 16.6]
Chronic musculoskeletal pain	10.2	[8.8 to 11.6]	5.4	[4.2 to 6.7]	7.4	[4.2 to 10.6]
Age^2	-	-	-	-	-1.2	[-2.4 to 0.0]
Age^3	-	-	-	-	0.1	[0.0 to 0.2]
Sex:Diabetes	-	-	-	-	-2.3	[-7.2 to 2.6]
Sex:Chronic respiratotry symptoms	-	-	-	-	-1.9	[-6.3 to 2.5]
Age:Hypertension	-	-	-	-	1.1	[0.2 to 2.1]
Age:Depression	-	-	-	-	-0.5	[-1.6 to 0.6]
Age:Chronic respiratory symptoms	-	-	-	-	-0.9	[-1.9 to 0.1]
Age:Chronic musculoskeletal pain	-	-	-	-	-0.4	[-1.1 to 0.2]

Coeff: regression coefficient/ CI: confidence interval
** In the case of dichotomous variables, regression coefficients can be interpreted as difference in means.*
*** Model 1 corresponds to the fully adjusted model: Disability = B_0 + B_1Sex + B_2Age + B3Education1 + B_4Education2 + B_5Depression + B_6Diabetes + B_7Hipertension + B_8Chronic respiratory symptoms + B_9Chronic musculoskeletal pain*
**** Model 2 corresponds to the fully adjusted model plus interaction terms: Disability = B_0 + B_1Sex + B_2Age + B_3Education1 + B_4Education2 + B_5Depression + B_6Diabetes + B_7Hipertension + B_8Chronic respiratory symptoms + B_9Chronic musculoskeletal pain + B_{10}Age^2 + B_{11}Age^3 + B_{12}Sex:Diabetes + B_{13}Sex:Chronic respiratory symptoms + B_{14}Age:Hypertension + B_{15}Age:Depression + B_{16}Age:Chronic respiratory symptoms + B_{17}Age:Chronic musculoskeletal pain*

Participants who were women, of older age, and who had lower levels of education showed higher disability in all models. Among diseases, depressive episodes were associated with the highest disability score in all models. According to model 1, people with depression showed a 12.1 [10.5 - 13.8] higher disability score (scale 0 - 100) than people without depression. Hypertension is the disease with the largest relative change in the disability score, comparing the unadjusted (model 0; regression coefficient 7.6 [5.9 - 9.3]) with the fully adjusted analysis (model 1; regression coefficient 1.5 [-0.1 to 3.1]). Regression coefficients from model 2 are more difficult to interpret because of the presence of interaction terms.

Table 3.2 summarises the estimates of attributable disability (DA) for each disease according to the population average-level approach, which are equivalent to the regression coefficients of the unadjusted models 0 and 1 (from Table 3.1). The inclusion of interaction terms for some diseases did not alter the estimates of DA, compared with those observed in the less complex models. DA can be interpreted as the excess of disability attributable to a certain health condition in people with that disease. For example, according to model 2, among people with a depressive episode, it is expected that 12.2 [10.4 - 14.1] units of their disability are attributable to this condition.

The second column of Table 3.2 presents the burden of disability due to each disease, estimated following a population average-level approach. Overall, the burden diminishes after adjustment by covariables, including comorbidities. Depressive episodes also show the highest burden of total population disability, based on the adjusted models (6.6% [5.5 – 7.7]). The burden of hypertension decreases from 6.5% in the bivariate analysis to 1.3%, after controlling for covariables, which highlights the importance of adjusting by covariables.

Estimates of the disability burden using an individual-level approach (third column of Table 3.2) are almost identical to those using the population average-level approach, although the former showed narrower uncertainty intervals. The burden of disability estimated, allowing the inclusion of significant interaction terms (model 2), does not differ from the simplest models. Results stratified by sex are available in the supplementary material (Tables S3.2 and S3.3, Figures S3.2 and S3.5).

Table 3.2. Attributable disability [0 – 100], and burden of disease [%] on disability for different diseases according to different models. Chilean National Health Survey 2009-2019 (n=4,447).

	Results obtained using parameters from Model 0					
	D_A (disability attributable)		Burden (population average-level approach)		Burden (individual-level approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depressive episode	15.6	[13.8 to 17.4]	8.5	[7.2 - 9.8]	8.5	[7.6 - 9.4]
Diabetes	11.0	[8.6 to 13.5]	2.3	[1.7 - 3.0]	2.3	[2.0 - 2.7]
Hypertension	7.6	[5.9 to 9.2]	6.5	[5.0 - 7.9]	6.5	[6.0 - 7.0]
Chronic respiratory symptoms	8.5	[5.8 to 11.2]	2.4	[1.6 - 3.3]	2.4	[2.0 - 2.8]
Chronic musculoskeletal pain	10.2	[8.8 to 11.6]	9.9	[8.5 - 11.4]	9.9	[9.2 - 10.7]

	Results obtained using parameters from Model 1					
	D_A (disability attributable)		Burden (population average-level approach)		Burden (individual-level approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depressive episode	12.1	[10.5 to 13.8]	6.6	[5.5 - 7.7]	6.6	[5.9 - 7.3]
Diabetes	4.4	[2.0 to 6.7]	0.9	[0.4 - 1.4]	0.9	[0.8 - 1.1]
Hypertension	1.5	[-0.1 to 3]	1.3	[-0.1 - 2.6]	1.3	[1.2 - 1.4]
Chronic respiratory symptoms	6.1	[3.7 to 8.5]	1.7	[1.0 - 2.5]	1.7	[1.4 - 2.0]
Chronic musculoskeletal pain	5.4	[4.2 to 6.7]	5.3	[4.1 - 6.5]	5.3	[4.9 - 5.7]

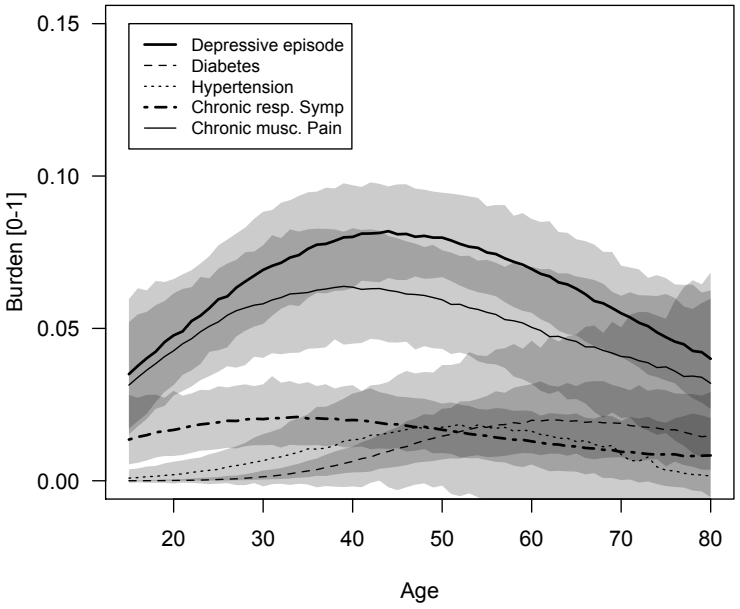
	Results obtained using parameters from Model 2					
	D_A (disability attributable)		Burden (population average-level approach)		Burden (individual-level approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depressive episode	12.2	[10.4 to 14.1]	6.6	[5.5 - 7.8]	6.6	[5.9 - 7.3]
Diabetes	4.5	[2.0 to 6.9]	0.9	[0.4 - 1.5]	0.9	[0.8 - 1.1]
Hypertension	1.2	[-0.6 to 3.0]	1.0	[-0.5 - 2.5]	1.3	[1.2 - 1.4]
Chronic respiratory symptoms	6.0	[3.4 to 8.6]	1.7	[0.9 - 2.5]	1.7	[1.4 - 2.0]
Chronic musculoskeletal symptoms	5.4	[3.9 to 6.9]	5.2	[3.8 - 6.7]	5.3	[4.9 - 5.7]

UI: uncertainty intervals (represent 2.5 and 97.5% quantile of resultant distribution)
D_A: the attributable disability shown is that estimated through the population average-level approach.

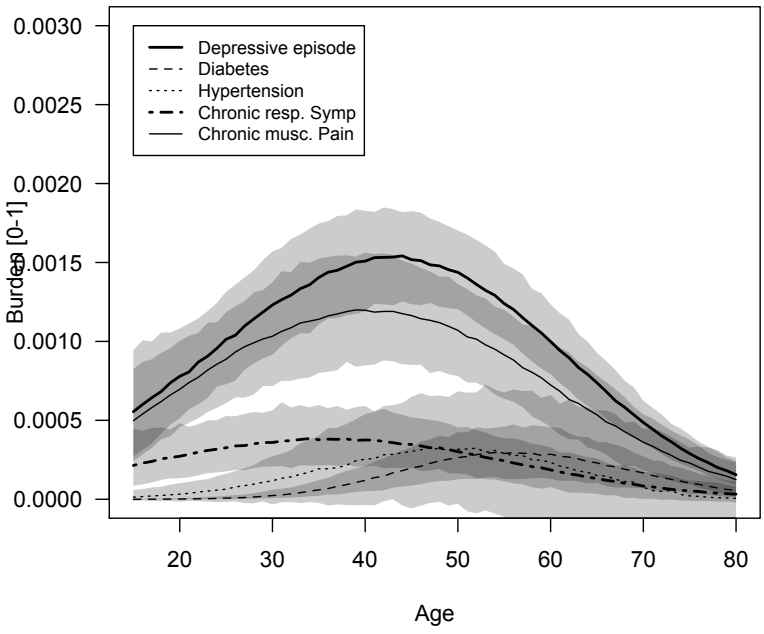
Figure 3.3, Panel A, presents the burden estimated for years of age for each disease under study, using model 1. At 40 years old, almost 8% of the disability burden is attributable to depressive episodes, followed by roughly 6% attributable to chronic musculoskeletal pain. Both conditions dominate the burden across all ages. The burden of hypertension and diabetes gain relevance at older ages. When the weight of each age is considered (Figure 3.3, Panel B), the largest contribution to the burden is for depressive episodes among people in their forties, which account for almost 0.15% of the total population disability.

Figure 3.3. Burden of disease attributable to five selected health condition according to age. Based on data from the Chilean National Health Survey 2009-2019 (n=4,447).

Panel A. Burden of disease attributable to five selected conditions according to age.



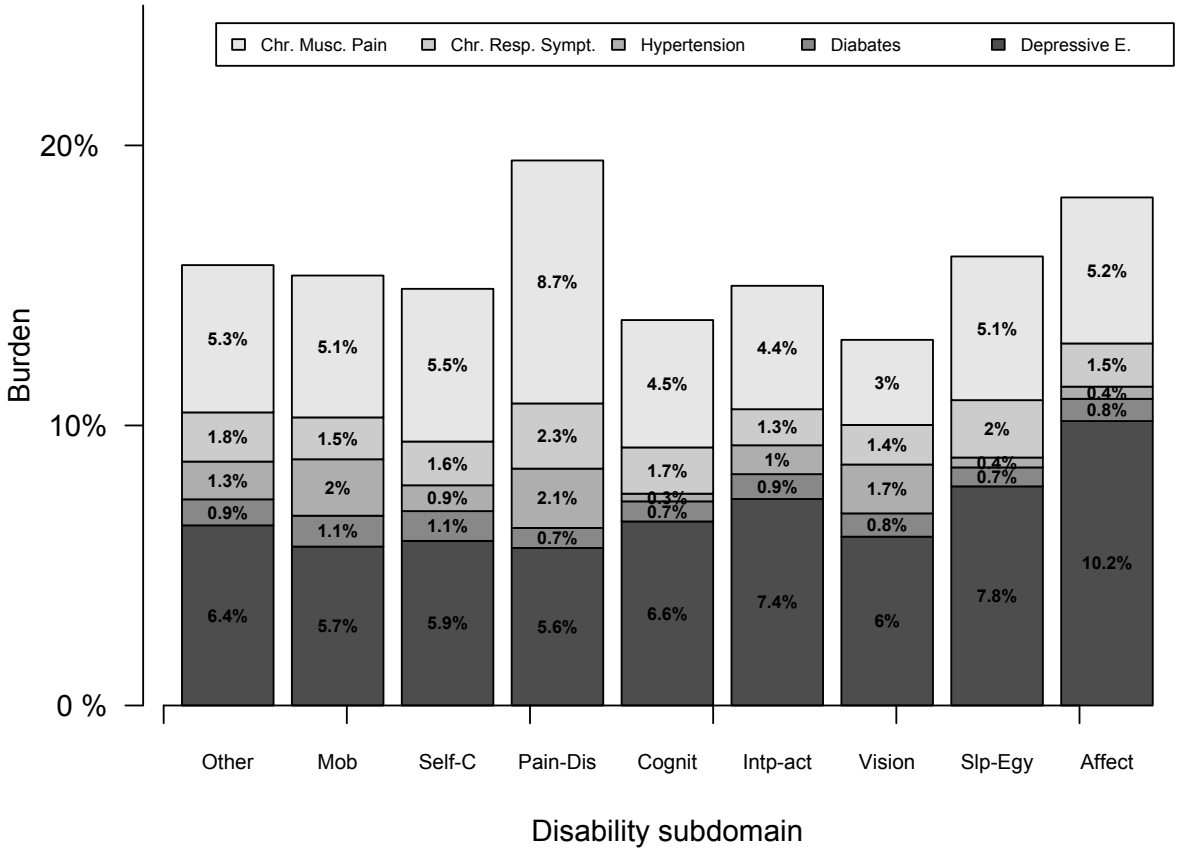
Panel B. Burden of disease attributable to five selected conditions according to age, adjusted by the weighted of disability at each age on the entire population disability.



Chronic resp. Symp: chronic respiratory symptoms/ Chronic musc. Pain: chronic musculoskeletal pain

The burden of disability due to diseases calculated using inputs from model 1 for each sub-domain of disability is presented in Figure 3.4. The diseases under study accounted for between 15% and 20% of the burden on disability for each sub-domain. Depressive episodes showed the greatest burden on disability for the sub-domains of affect, sleep and energy, and interpersonal activities. The burden attributable to diabetes was small and homogeneous across different sub-domains, whilst, in the case of hypertension, the sub-domains of pain and discomfort, mobility and vision were more prominent. The burden due to chronic respiratory symptoms was relatively homogeneous across sub-domains, while chronic musculoskeletal pain showed its greatest burden in the sub-domain of pain and discomfort. Results, according to sex, are available in the supplementary material (Figures S3.3, S3.4, S3.6, and S3.7).

Figure 3.4. Burden of disease on sub-domains of disability for five non-communicable diseases. Chilean National Health Survey 2009-2019 (n=4,447).



Chr. Resp. Sympt. : chronic respiratory symptoms/ Chr. Musc. Sympt: chronic musculoskeletal symptoms.
Sub-domain of disability: Affect, Slp-Egy (sleep and energy), Vision, Intp-act (interpersonal activities), Cognit (cognition), Pain-Dis (pain and discomfort), Self-C (self-care), Mob (mobility), Other (overall health)

Finally, the fraction of disability attributable to each sub-domain according to disease, standardised by the same profile of the population in terms of covariables, is presented in the supplementary material (Figure S.3.1). It is shown that the AF differs little across diseases.

3.5. Discussion

In the background, I mentioned that I would gather two metrics in the proposal to calculate the burden of disability attributable to diseases: the burden of disease and the AF. I also mentioned that the burden of disability is traditionally measured through the YLDs, whilst the AF is used to estimate the impact of risk factors on diseases.

In the procedure developed in this chapter, I am following an approach from AFs. However, instead of using a health condition as the outcome, disability is used, and instead of using a risk factor as an exposure, the presence of a disease is used.

In the methods and results part of this chapter, I have shown the properties of the proposal following this framework. I showed that the procedure seems to be a flexible and practical way to estimate the burden of disability due to diseases using AF applied on a continuous outcome.

This new procedure showed similar results through different approaches (population average-level and individual-level), initial specifications (model 1, model 2), and, in comparison with other methodological approaches such as the YLD, it seems more flexible and simpler to implement. Also, I showed that it is possible to apply the same procedure to explore the contribution of each health condition to different sub-domains of disability, as well as the contribution of these sub-domains to the overall disability attributable to a specific disease.

3.5.1. The procedure addresses two challenges of AF metrics

The AF family of effect measures have been historically developed for dichotomous outcomes, mainly applied to data arranged in a contingency table, and more recently applied through logistic regression models [42, 43, 63-65]. Here, I show how to apply AF on a continuous outcome, in a way that maintains consistency with the original formulation provided by Levin et al. in 1953 [31, 33, 66, 67]. Unlike traditional formulations of AF, here it uses neither contrasted incidence rates, nor incidence risks, nor prevalent cases between different strata, but different expected values and distributions of quantities, such as the

disability score. The procedure implicitly assumes that these quantities from different individuals are summable, which is an assumption shared with other metrics such as YLDs [3].

Some studies [68] have documented multivariable approaches such as the one presented by Tanusepruto et al., which was introduced in the background section, and the one I am presenting here, which produces more stable estimates of AF than other traditional approaches [42].

On the other hand, very few studies have used AF directly to estimate the burden of disability due to diseases, and all of them have considered the outcome variable – i.e. disability – as a dichotomous variable (disabled or not disabled). However, assuming disability as a binary construct goes against the current conceptual frameworks of disability [69], and it is inefficient as it loses information [70]. Moreover, as far as I am aware, only Palazzo et al. studied the burden of disease by applying AF to disability, which also included an analysis by sub-domains [49].

3.5.2. Properties of the burden of disability attributable to diseases

In the methods section, it was mentioned that the burden of disability attributable to diseases is equivalent to the concept of PAF. Consequently, and perhaps more accurately, this metric could be called the 'fraction of the population disability attributable to diseases'. However, for convenience, I prefer to use the term burden of disability, as it refers to a metric already known, which makes it easier to understand. Moreover, this metric can be expressed not only as a fraction of the total population disability, but also as an absolute measure, which can be used to calculate rates.

As with any set of PAF calculated following the framework introduced in the background, the sum of the burdens from different diseases may add up more than 100% (see section 3.2.2.8). This is because of the potential confusion between diseases, and also because a disease can play an intermediate role in the pathway from another disease generating disability [46]. Both elements can be resolved through an adequate adjustment for the presence of comorbidity in the regression models used to estimate the PAF. However, it is necessary to be cautious when intermediation is suspected, since the burden calculated from the adjusted

models will not include the effect of the factors that are downstream to the disease under study. For example, the burden attributable to musculoskeletal chronic pain calculated by adjusting depressive episodes does not consider the potential benefit that removing this disease may cause the burden due to depression.

How much this factor affects the burden, when estimated for a specific disease, can be observed in the extent of the change in the burden after removing from the regression models the comorbidities that could be plying an intermediate role.

There are other alternatives to overcome the challenge of estimating the PAF within the limits of the unit (100%). One of them is the 'average AF', which is a procedure based on averaging multiple PAFs, calculated by removing each health condition (or risk factor) progressively and randomly [71]. Another approximation is using 'attributional methods', which are based on similar procedures applied for decomposing the life expectancy by different causes of mortality [72]. However, some studies have shown that the results obtained by different procedures are rather equivalent [48].

The fraction of the burden of disability attributable to a disease using YLD does not have the limitation of potentially surpassing 100%, since the calculus is restricted to the disorders in which the burden was estimated, a set that is assumed to be exhaustive. However, similar to the approach suggested in this thesis, the YLD has the potential limitation of overestimating the disability due to insufficient adjustment by potential confounders (see chapter two). Moreover, YLD neither consider potential intermediations, nor interactions between diseases producing disability.

Another important feature of the burden of disability attributable to diseases is that, similar to any other PAF, we are studying the contrast between a current scenario and a hypothetical counterfactual scenario, where the health condition was removed entirely. However, in other contexts, intermediate counterfactual scenarios can be implemented, as in the case of effective coverage estimation.

Finally, as in the case of PAF calculated for risk factors, the burden of disability attributable to diseases is the lowest limit of the etiological fraction of the disability attributable to a health condition. This means that, through the approaches presented in this chapter, we can measure the disability in excess attributable to a specific disease. In other words, we are

calculating the averted disability that will not be replaced by another source of disability, if we remove a particular health condition.

3.5.3. Difference between approaches

The two approaches I explored to estimate the burden of disability – that is the population average-level and the individual level – yielded no substantial differences in the results, although the latter produced narrower intervals of uncertainty. If both approaches generate an equivalent result, it may be asked why we do not simply choose one of them.

The answer is twofold. Firstly, because the intermediate results from both approaches are different and separately provide valuable information. For example, the D_A from the population average-level approach corresponds to the average disability attributable to the disease at the individual level. On the contrary, the D_A from the individual-level approach corresponds to the total units of disability attributable to the disease in the entire population, and this information can be used to calculate rates of disabilities. For example, Table 3.3 presented depressive episode estimates for D_0 (disability from people without the disease), D_1 (disability from people with the disease), and D_A (the disability attributable to the disease) using the regression model 1 as the initial specification, in accordance with both approaches.

In that table, we can observe that the disability attributable to depression calculated using the population average-level approach is 12.1 [10.5 – 13.7], which corresponds to the average disability in excess that is attributable to the disease in a person with that health condition. On the other hand, D_1 (44.9 [43.5 – 46.3]) is the total disability expected in a person with depression, and D_0 (29.3 [28.6 – 29.9]) is the expected disability in a person without depression. The division between D_A and $D_0 + D_1$, weighted by the prevalence of the disorder, produces the estimate of the burden (6.6% [5.5 – 7.7]).

Table 3.3. Different outputs from the process to estimate the burden of disability attributable to diseases, according to population average-level and individual-level approaches, and comparison with estimates from the Institute of Health Metric and Evaluation (IHME) for depressive episode in Chile 2010.

	Population average-level approach			Individual-level approach*		
	Estimate	**LCI**	**UCI**	**Estimate**	**LCI**	**UCI**
Burden	6.6%	5.5%	7.7%	6.6%	5.9%	7.3%
D_A	12.1	10.5	13.7	231,541	209,928	253,385
D_0	29.3	28.6	29.9	2,648,819	2,535,358	2,762,280
D_1	44.9	43.5	46.3	856,724	775,820	937,629

	Individual-level approach (rates x 100,000 Inhabitants) *			Burden and YLD for Major Depressive Disorder according to IHME for Chile, 2010‡		
	Rate (x 100.100)	**LCI**	**UCI**	**Estimate**	**LCI**	**UCI**
Burden	-	-	-	5.8%	4.7%	7.0%
D_A (YLD)	2,114.0	1,916.6	2,313.4	113,136	79,326	152,997
D_0	24,183.7	23,147.8	2,5219.6	-	-	-
D_1	7,821.9	7,083.2	8,560.5	-	-	-

Estimates using data from the Chilean National Health Survey 2009-2019 (n=4,447)
D_A: disability attributable to depression. / D_0: disability among people without depression. / D_1: disability among people with depression/ YLD: years lost by disability/ LCI: lower creditable interval/ UCI: upper creditable interval.
* *For reasons of comparability, results from the population average-level approach were divided by 100, which rescaled the score of disability to a range between 0 and 1, instead 0 and 100.*
‡ *Extracted data from data https://vizhub.healthdata.org/gbd-compare/ (consulted in September, 2020)*

On the other hand, the estimate for DA using the individual-level approach, after rescaling the score of disability to a range between 0 and 1, is 231,541 [209,928 – 253,385], which is equivalent to 231,541 people with extreme disability (i.e. score = 1). Similarly, the total disability from people with depression (D_1) is 856,724 [775,820 – 937,629], and from people without depression (D_0) is 2,648,819 [2,535,358 – 2,762,280]. Dividing D_A by $D_0 + D_1$, using these quantities, gives a burden of 6.6% again. In addition, the amount of disability attributable (D_A) can be transformed into a rate, dividing it by the total number of people exposed to have a depressive episode, which is equivalent to the total population represented by the survey (10,952,907 inhabitants). That results in 2,114 cases of extreme disability attributable to depressive disorder by 100,000 inhabitants. This expression allows easy comparison of the burden of disability attributable to diseases in absolute terms, across different conditions and populations.

Moreover, if we assume that the prevalence of depression remains constant during a year and that the population is fixed, we can express the estimate of D_A as 231,541 person-years of extreme disability. This number can be compared with the YLD estimated for depression by the IHME for Chile, in the year 2010, which is 113,136 (see Table 3.3). Two main elements can explain differences between our estimates and those from the IHME.

Firstly, the disability weights that we are using are not the same as those used by the IHME. It is worth noting that the D_A, estimated through the population average-level approach, is equivalent to the result of a health state valuation and can be interpreted as disability weights (see the discussion of chapter two). Moreover, our measure of disability weights is anchored in the concepts of full health and extreme disability, while the measure for the YLD is anchored in full health and death. Despite these differences, I showed in chapter two, that the disability weights that we are using are quite similar to those from the IHME.

Secondly, and more importantly, the prevalence of depressive episodes in the past 12 months used in this study was 17.4% [15.5-19.3], which is high compared to international estimates – usually around to 5-10% [73]. It is likely that the IHME used an estimate for the prevalence of depression in Chile lower than the actual observed using data from the Ch-NHS.

Despite the differences in the number of person-years of extreme disability attributable to depression from our study, and the total YLD by depression according to the IHME, when we compare the relative estimates, we do not find much difference. The IHME estimated that 5.8% [4.7 – 7.0] of the total YLDs were attributable to depression in Chile, 2010, with an uncertainty that overlaps with the estimate presented in this chapter (6.6% [5.9 – 7.3]). Notice that, to achieve this fraction, the IHME first had to estimate the YLD for 359 other health conditions.

The second important reason to take into consideration for both approaches (individual-level and population average-level) is related to their computational implementation. In general terms, the individual-level approach is more straightforward than implemented based on the population average-level (see and compare 'Main functions for R S3.1' in the supplementary material). However, when it is required to describe estimates against a continuous variable (e.g. age in Figure 3.3), the population average-level is more convenient.

To conclude, both approaches are complementary, and both provide different but valuable information about the burden of disability attributable to diseases, even though the final estimates of the burden are similar.

3.5.4. Difference between initial specifications

In the case of the diseases in this study, we observed important differences in the burden before and after adjustment by covariates, including comorbidities. However, these differences were not observed between the fully adjusted model (model 1) and the fully adjusted model considering interactions (model 2).

The reduction of the burden after adjustment is most likely due to the confounding effect of covariates, including comorbidities. Besides, the absence of a substantial change in the burden using models with and without interaction terms shows that this specification, in practice, does not add much to the overall estimates of disability, at least using the present data.

Also, it shows that both approaches (i.e. population average-level and individual-level) are flexible enough to produce results with different levels of complexity in their assumptions, while still remaining consistent. Exploring interaction terms in this study was also justified, because it has been suggested that the experience of disability can differ substantially according to sex or age in people with chronic conditions [74, 75].

3.5.5. Usefulness of sub-domains of disability

Using the variables for sub-domains of disability, as described in chapter two, made it possible to estimate the burden of diseases for each sub-domain. Also, it allowed the decomposing of the burden in terms of 'activities and participations' affected by each disease being studied.

This kind of analysis can be useful for a better understanding of the pathways and mechanisms involved in generating disability within each disease. Interestingly, the fraction attributable to each sub-domain was similar across the diseases (Figure S.3.1), which may

reflect the fact that the distribution of disability by sub-domains is more dependent on structural elements underlying the health conditions. That was not the case for the analysis of the burden by sub-domain (Figure 3.4), where the contribution by each disease changed significantly in a way more or less expected.

Regardless of the specific conclusions that we can achieve from the data, the procedure introduced to estimate the burden of disability by sub-domain attributable to diseases shows an exciting opportunity for further analysis of disability and its relationship whit diseases. What's more, in the context of effective coverage, it will allow the impact of specific components of healthcare intervention to be explored in more detail (see chapter six for further development).

3.5.6. Advantages of the procedure

The two approaches outlined have several strengths in comparison to other metrics of disability, such as YLD. Firstly, we used the same unique and local source of information to obtain figures for the prevalence and disability of diseases, ensuring consistency between parameters and the target population. Moreover, throughout our procedure, the disability weights (i.e., D_A) are estimated in only one step that integrates: i) the identification of all health states associated with a particular disease, ii) their relative frequency, and also iii) the specific disability-related to each health state (see chapter two).

Secondly, the procedure is based on a linear regression model, which is familiar to most epidemiologists and health scientists, allowing enough flexibility to adjust estimates by comorbidity, to introduce interaction terms, or to do both when this is deemed necessary. Moreover, the procedure to adjust disability weights (i.e., D_A) is much simpler to implement than the one used by the IHME, which is based on multiple steps and sources of information, as well as enormous amounts of simulations for each stratum of sex and age [19].

Thirdly, the whole procedure can be easily packaged as a function for statistical software, because it requires only one source of information and is based on few steps. See 'Main functions for R S3.1' in the supplementary material.

Fourthly, the approach allows us to estimate the relative weight of a small number of diseases over the total population disability, without the need to analyse an extensive list of health conditions (the GBD first needs to calculate the burden for more than 350 diseases).

Finally, and perhaps most importantly, the procedure works similarly well with an indicator like disability as it does with a continuous indicator of health status, such as health related to quality of life, or health state utilities [76], offering a more comprehensive range of outcome options. For example, using the complement of health state utilities (i.e. 1 – health state utilities) and following the individual-level approach, the disability attributable would correspond to the total health state utilities lost in a year due to the disease under study. This result could be assimilated to an equivalent concept of 'prevalent' Quality-Adjusted Life Years (QALY), a metric often used in the field of health technology assessment [77]. Furthermore, D_A estimated through the population average-level approach would correspond to the adjusted difference in health state utilities between those with and without a disease. That is a piece of valuable information required in models for cost-effective analysis, which commonly has to be estimated through small samples of patients or expert elicitations. Similarly, D_0 can be used as a basal health state; another input generally required in that type of analysis [78]. Moreover, since health state utilities are anchored in full health and death, the results are easier to combine with mortality data. Recently, I published a piece of research where, using the approach proposed in this chapter, I compared results using 'prevalent' QALYs versus YLD [79]. In another research paper, also recently published, I presented a preliminary proposal for the procedure used in this chapter [80].

3.5.7. Limitations

The main limitation of the approach presented in this chapter is related to the assumption of causality between exposure to diseases and attributable disability obtained from cross-sectional data [81]. Although the procedure allows for the models to be adjusted according to comorbidities and other potential confounders, as well as the inclusion of interaction terms, it is not possible to rule out residual confounding completely.

A second limitation is that the number of diseases whose burden can be evaluated is restricted to those included in the design of the surveys. Health conditions with a small

prevalence, or those of a short duration, are usually excluded from this kind of study [44]. Therefore, this approach could be especially suitable for chronic non-communicable diseases. A further discussion of limitations is developed in chapter six.

3.6. Conclusion

The procedure introduced here, based on AF applied to a continuous outcome, may provide a simpler alternative to monitoring the burden of disability due to diseases, especially for chronic and non-communicable conditions, in contexts where national health surveys are available. This approach can be further explored using outcomes other than disability, such as health-related quality of life, or health state utilities, which can be suitable to guide the decision-making process.

The procedure also allows us to explore the burden of disability by sub-domains across diseases, which can offer additional information for policymakers to address more accurately the health and social needs of populations, informing new targets and programmes.

Finally, the procedure is a central element used to estimate effective coverage, which will be addressed in more detail in the following chapter.

References of chapter three

1. Shengelia B, Tandon A, B. Adams O, Murray CJL: **Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy**. *Social Science & Medicine* 2005, **61**:97-109.
2. Bank W: **World Development Report 1993. Investing in Health**: Oxford University Press; 1993.
3. Murray CJ, Lopez AD: **Quantifying disability: data, methods and results**. *Bull World Health Organ* 1994, **72**(3):481-494.
4. Murray CJ, Lopez AD: **Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study**. *Lancet* 1997, **349**(9063):1436-1442.
5. Bank TIBfRaDTW: **Global Burden of Disease and Risk Factors**: Oxford University Press and The World Bank; 2006.
6. World Health Organization: **The global burden of disease: 2004 update.**: World Health Organization 2008.
7. Group PH: **Victorian Burden of Disease Study. Mortality and morbidity in 2001.**: Public Health Group, Rural and Regional Health and Aged Care Services Division. Victorian Government Department of Human Services. Melbourne Victoria; 2005.
8. Chile MdSd: **Estudio de Carga de Enfermedad en Chile**. In.; 1996.
9. Chile MdSd, Pontifica Universidad Católica de Chile DdSP: **Informe Final Estudio de Carga de Enfermedad y Carga Atribuible**. In.: Ministerio de Salud de Chile; 2008.
10. Mathers C, Vos T, Lopez A, J. S, M E: **National Burden of Disease Studies: A Practical Guide. Edition 2.0.**: Global Program on Evidence for Health Policy. World Health Organization.; 2001.
11. Moon L, Gourley M, Goss J, Lum On M, Laws P, Reynolds A, Juckes R: **History and development of national burden of disease assessment in Australia**. *Archives of Public Health* 2020, **78**(1):88.
12. Hilderink HBM, Plasmans MHD, Poos M, Eysink PED, Gijsen R: **Dutch DALYs, current and future burden of disease in the Netherlands**. *Arch Public Health* 2020, **78**:85.
13. World Health Organization: **The World Health Report 2000. Health Systems: Improving Performance.** : World Health Organization; 2000.
14. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S *et al*: **Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010**. *Lancet* 2012, **380**(9859):2197-2223.
15. Murray CJ, Barber RM, Foreman KJ, Abbasoglu Ozgoren A, Abd-Allah F, Abera SF, Aboyans V, Abraham JP, Abubakar I, Abu-Raddad LJ *et al*: **Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: quantifying the epidemiological transition**. *Lancet* 2015, **386**(10009):2145-2191.
16. GBD: **Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a**

systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016, **388**(10053):1603-1658.

17. DALYs GBD, Collaborators H: **Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017**. *Lancet* 2018, **392**(10159):1859-1922.
18. Murray CJ: **Quantifying the burden of disease: the technical basis for disability-adjusted life years**. *Bull World Health Organ* 1994, **72**(3):429-445.
19. Murray CJ, Ezzati M, Flaxman AD, Lim S, Lozano R, Michaud C, Naghavi M, Salomon JA, Shibuya K, Vos T *et al*: **GBD 2010: design, definitions, and metrics**. *Lancet* 2012, **380**(9859):2063-2066.
20. Salomon JA, Haagsma JA, Davis A, de Noordhout CM, Polinder S, Havelaar AH, Cassini A, Devleesschauwer B, Kretzschmar M, Speybroeck N *et al*: **Disability weights for the Global Burden of Disease 2013 study**. *The Lancet Global health* 2015, **3**(11):e712-723.
21. Disease GBD, Injury I, Prevalence C: **Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017**. *Lancet* 2018, **392**(10159):1789-1858.
22. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, Begum N, Shah R, Karyana M, Kosen S *et al*: **Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010**. *Lancet* 2012, **380**(9859):2129-2143.
23. Haagsma JA, Maertens de Noordhout C, Polinder S, Vos T, Havelaar AH, Cassini A, Devleesschauwer B, Kretzschmar ME, Speybroeck N, Salomon JA: **Assessing disability weights based on the responses of 30,660 people from four European countries**. *Population health metrics* 2015, **13**:10.
24. Mont D: **Measuring health and disability**. *Lancet* 2007, **369**(9573):1658-1663.
25. Reidpath DD, Allotey PA, Kouame A, Cummins RA: **Measuring health in a vacuum: examining the disability weight of the DALY**. *Health Policy Plan* 2003, **18**(4):351-356.
26. Burstein R, Fleming T, Haagsma J, Salomon JA, Vos T, Murray CJ: **Estimating distributions of health state severity for the global burden of disease study**. *Population health metrics* 2015, **13**:31.
27. Voigt K, King NB: **Disability weights in the global burden of disease 2010 study: two steps forward, one step back?** *Bull World Health Organ* 2014, **92**(3):226-228.
28. Nord E: **Uncertainties about disability weights for the Global Burden of Disease study**. *The Lancet Global health* 2015, **3**(11):e661-662.
29. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V *et al*: **Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010**. *Lancet* 2012, **380**(9859):2163-2196.

30. Shiffman J, Shawar YR: **Strengthening accountability of the global health metrics enterprise**. *Lancet (London, England)* 2020, **395**(10234):1452-1456.
31. Poole C: **A history of the population attributable fraction and related measures**. *Ann Epidemiol* 2015, **25**(3):147-154.
32. Cornfield J: **A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix**. *J Natl Cancer Inst* 1951, **11**(6):1269-1275.
33. Leviton A: **Letter: Definitions of attributable risk**. *Am J Epidemiol* 1973, **98**(3):231.
34. MacMahon B, Pugh TF, Ipsen J: **Epidemiologie Methods**. *Epidemiologie Methods* 1960.
35. Cole P, MacMahon B: **Attributable risk percent in case-control studies**. *Br J Prev Soc Med* 1971, **25**(4):242-244.
36. Lilienfeld AM: **Epidemiology of infectious and non-infectious disease: some comparisons**. *Am J Epidemiol* 1973, **97**(3):135-147.
37. Miettinen OS: **Proportion of disease caused or prevented by a given exposure, trait or intervention**. *Am J Epidemiol* 1974, **99**(5):325-332.
38. Walter SD: **The estimation and interpretation of attributable risk in health research**. *Biometrics* 1976, **32**(4):829-849.
39. Ouellet RP, Apostolides AY, Entwisle G, Hebel JR: **Estimated community impact of hypertension control in a high risk population**. *Am J Epidemiol* 1979, **109**(5):531-538.
40. Morgenstern H, Bursic ES: **A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population**. *J Community Health* 1982, **7**(4):292-309.
41. Murray CJ, Lopez AD: **On the comparable quantification of health risks: lessons from the Global Burden of Disease Study**. *Epidemiology* 1999, **10**(5):594-605.
42. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C: **Estimating the population attributable risk for multiple risk factors using case-control data**. *Am J Epidemiol* 1985, **122**(5):904-914.
43. Greenland S, Drescher K: **Maximum likelihood estimation of the attributable fraction from logistic models**. *Biometrics* 1993, **49**(3):865-872.
44. KJ R, S G, TL L: **Modern epidemiology**. Philadelphia: Lippincott Williams & Wilkins; 2008.
45. Greenland S, Robins JM: **Conceptual problems in the definition and interpretation of attributable fractions**. *Am J Epidemiol* 1988, **128**(6):1185-1197.
46. World Health Organization: **Comparative Quantification of Health Risks. Global and Regional Burden of Disease Attributable to Selected Major Risk Factors**: World Health Organization; 2004.
47. Tanuseputro P, Perez R, Rosella L, Wilson K, Bennett C, Tuna M, Hennessy D, Manson H, Manuel D: **Improving the estimation of the burden of risk factors: an illustrative comparison of methods to measure smoking-attributable mortality**. *Population health metrics* 2015, **13**:5.
48. Palazzo C, Yokota RTC, Ferguson J, Tafforeau J, Ravaud JF, Van Oyen H, Nusselder WJ: **Methods to assess the contribution of diseases to disability**

using cross-sectional studies: comparison of different versions of the attributable fraction and the attribution method. *Int J Epidemiol* 2018.

49. Palazzo C, Ravaud JF, Papelard A, Ravaud P, Poiraudeau S: **The burden of musculoskeletal conditions**. *PLoS One* 2014, **9**(3):e90633.
50. Salomon JA, Mathers CD, Chatterji S, Sadana R, Üstün TB, Murray CJL: **Quantifying Individual Levels of Heath: Definitions, Concepts, and MEasurement Issues**. In: *Health Systems Performance Assessment Debates, Methods and Empiricism.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 301-318.
51. Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJ: **Selected major risk factors and global and regional burden of disease**. *Lancet* 2002, **360**(9343):1347-1360.
52. Ezzati M, Hoorn SV, Rodgers A, Lopez AD, Mathers CD, Murray CJ: **Estimates of global and regional potential health gains from reducing multiple major risk factors**. *Lancet* 2003, **362**(9380):271-280.
53. Rodgers A, Ezzati M, Vander Hoorn S, Lopez AD, Lin RB, Murray CJ: **Distribution of major health risks: findings from the Global Burden of Disease study**. *PLoS Med* 2004, **1**(1):e27.
54. World Health Organization: **Global health risks: mortality and burden of disease attributable to selected major risks.**: World Health Organization 2008.
55. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Amann M, Anderson HR, Andrews KG, Aryee M *et al*: **A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010**. *Lancet* 2012, **380**(9859):2224-2260.
56. GBD: **Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015**. *Lancet* 2016, **388**(10053):1659-1724.
57. GBD: **Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016**. *Lancet* 2017, **390**(10100):1345-1422.
58. GBD: **Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017**. *Lancet* 2018, **392**(10159):1923-1994.
59. Menezes AM, Perez-Padilla R, Jardim JR, Muino A, Lopez MV, Valdivia G, Montes de Oca M, Talamo C, Hallal PC, Victora CG *et al*: **Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study**. *Lancet* 2005, **366**(9500):1875-1881.
60. Bennett K, Cardiel MH, Ferraz MB, Riedemann P, Goldsmith CH, Tugwell P: **Community screening for rheumatic disorder: cross cultural adaptation and screening characteristics of the COPCORD Core Questionnaire in Brazil, Chile, and Mexico. The PANLAR-COPCORD Working Group. Pan American**

League of Associations for Rheumatology. Community Oriented Programme for the Control of Rheumatic Disease. *J Rheumatol* 1997, **24**(1):160-168.
61. Greenland S, Daniel R, Pearce N: **Outcome modelling strategies in epidemiology: traditional methods and basic alternatives**. *Int J Epidemiol* 2016, **45**(2):565-575.
62. King G, Tomz M, Wittenberg J: **Making the Most of Statistical Analyses: Improving Interpretation and Presentation**. *American Journal of Political Science* 2000, **44**(2).
63. Benichou J, Gail MH: **Variance calculations and confidence intervals for estimates of the attributable risk based on logistic models**. *Biometrics* 1990, **46**(4):991-1003.
64. Deubner DC, Wilkinson WE, Helms MJ, Tyroler HA, Hames CG: **Logistic model estimation of death attributable to risk factors for cardiovascular disease in Evans County, Georgia**. *Am J Epidemiol* 1980, **112**(1):135-143.
65. Drescher K, Schill W: **Attributable risk estimation from case-control data via logistic regression**. *Biometrics* 1991, **47**(4):1247-1256.
66. Levin ML: **The occurrence of lung cancer in man**. *Acta Unio Int Contra Cancrum* 1953, **9**(3):531-541.
67. WALTER SD: **The distribution of Levin's measure of attributable risk**. *Biometrika* 1975, **62**(2):371-372.
68. Tuson M, Turlach B, Vickery A, Whyatt D: **Reducing Bruzzi's Formula to Remove Instability in the Estimation of Population Attributable Fraction for Health Outcomes**. *American Journal of Epidemiology* 2017, **187**(1):170-179.
69. World Health Organization: **International Classification of Functioning, Disability and Health (ICF)**: World Health Organization; 2001.
70. Royston P, Altman DG, Sauerbrei W: **Dichotomizing continuous predictors in multiple regression: a bad idea**. *Stat Med* 2006, **25**(1):127-141.
71. Eide GE, Gefeller O: **Sequential and average attributable fractions as aids in the selection of preventive strategies**. *Journal of Clinical Epidemiology* 1995, **48**(5):645-655.
72. Nusselder WJ, Looman CW: **Decomposition of differences in health expectancy by cause**. *Demography* 2004, **41**(2):315-334.
73. Rai D, Zitko P, Jones K, Lynch J, Araya R: **Country- and individual-level socioeconomic determinants of depression: multilevel cross-national comparison**. *Br J Psychiatry* 2013, **202**(3):195-203.
74. Waterhouse P, van der Wielen N, Banda PC, Channon AA: **The impact of multi-morbidity on disability among older adults in South Africa: do hypertension and socio-demographic characteristics matter?** *Int J Equity Health* 2017, **16**(1):62.
75. Holman D: **Chronic conditions as predictors of later life disability employment exit: a gendered analysis**. *Occup Environ Med* 2019, **76**(7):441-447.
76. Salomon JA, Murray CJ: **A multi-method approach to measuring health-state valuations**. *Health Econ* 2004, **13**(3):281-290.
77. Ara R, Wailoo A: **Using health state utility values in models exploring the cost-effectiveness of health technologies**. *Value Health* 2012, **15**(6):971-974.

78. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, GW T: **Methods for the Economic Evaluation of Health Care Programmes**, 4th edn. Oxford, UK: Oxford University Press; 2015.
79. Zitko P, Bilbeny N, Balmaceda C, Abbot T, Carcamo C, Espinoza M: **Prevalence, Burden of Disease, and Lost in Health State Utilities attributable to Chronic Musculoskeletal Disorders and Pain in Chile**. *BMC Public Health* 2020, **(under second review)**.
80. Zitko P, Bilbeny N, Vargas C, Balmaceda C, Eberhard ME, Ahumada M, Rodríguez MF, Flores J, Markkula N, Espinoza MA: **Different Alternatives to Assess the Burden of Disease Using Attributable Fraction on a Disability Variable: The Case of Pain and Chronic Musculoskeletal Disorders in Chile**. *Value Health Reg Issues* 2021, **26**:15-23.
81. Flanders WD, Klein M, Mirabelli MC: **Conditions for valid estimation of causal effects on prevalence in cross-sectional and other studies**. *Ann Epidemiol* 2016, **26**(6):389-394.

Chapter 4: Proposal for an alternative way of calculating the effective coverage of a healthcare intervention

4.1. Introduction

In the first chapter, I presented the most widely accepted and used definition of effective coverage [1-4]. However, through a systematic scope review, I showed that almost none of the scientific articles that explored effective coverage use this approach. Moreover, the international organisations that currently track Universal Health Coverage goals globally also failed to follow this approach. The main reason was that they did not include health gains in the quality parameter of effective coverage.

To better understand the idea of health gains, in the second chapter I introduced basic concepts such as health, disability, health states, and health-disability-utility weights. In addition, I showed how, through a latent variable approach, it was possible to calculate disability weights associated with diseases, without the need to identify specific health states.

In the third chapter, I presented a practical approximation for calculating the burden of disability attributable to diseases, adapting the procedure used to estimate disability weights which was introduced in chapter two. This procedure was based on the concept of attributable fractions, which is a critical component of the proposal I will present for calculating effective coverage.

In this chapter, I will develop a formal proposal to estimate effective coverage for healthcare interventions, including health gains in the quality parameter, as suggested by Shengelia et al. [1]. As in the previous chapters, this procedure was developed using data from a national health survey.

In the second section of this chapter, I will present two general and complementary approaches for measuring effective coverage and other related concepts, based on the concept of health gains. In the third section, I will show how to implement such approaches, and in the fourth section, I will introduce the data and statistical procedures used to test these approaches. Then, in the fifth section, I will present the results of this test, while in the sixth and final section I will discuss the assumptions, limitations (including the possibility of reverse causation estimating health gain), and strengths of these approaches.

4.2. A practical proposal to measure effective coverage, including health gains

According to Shengelia et al., to calculate effective coverage we need to consider three elements: the need for a healthcare intervention; the utilisation of that healthcare intervention by the people that need it; and the quality of such an intervention. The latter, according to Shengelia et al., corresponds to the fraction between the current health gain of those with the need receiving the intervention, and the maximum possible health gain they could expect to receive [1] (see section 1.3 from chapter one).

It is important to remember that Shengelia et al. originally defined effective coverage at the individual level. Therefore, to obtain an overall indicator of effective coverage valid at the population level, the authors suggest aggregating the individual values of effective coverage and weighting them by the magnitude of the health gain. At a population level, effective coverage can also be understood as the product of the coverage (i.e. the fraction of people in need of a healthcare intervention that are using it), and the quality of that healthcare intervention (see chapter one for details).

In this thesis, I propose to express effective coverage as a function of the burden of disability attributable to a normative need for a healthcare intervention, using the concepts developed in chapters two and three. By normative need, I refer to those health states that could benefit from a healthcare intervention. From a normative perspective, the 'need' can assimilate the disease targeted by the healthcare intervention of interest [5]. For example, if we are interested in investigating the effective coverage of the treatment for a depressive disorder, we can consider people with a depressive disorder as those in need of such treatment. The underlying assumptions will depend on how a healthcare intervention is defined, which will be discussed later in this chapter. From here on, for the sake of simplicity, I will refer to the need for a healthcare intervention as a normative need.

Following the notion of the burden of disability attributable to disease, I propose to define effective coverage at the population level as: the fraction of the avoidable disability, attributable to a normative need, that is avoided by its corresponding healthcare intervention. Unlike Shengelia et al., this definition will be developed primarily at a population level. However, we can translate this definition to the individual level by changing the concept of the burden of disability attributable to the disability attributable. The

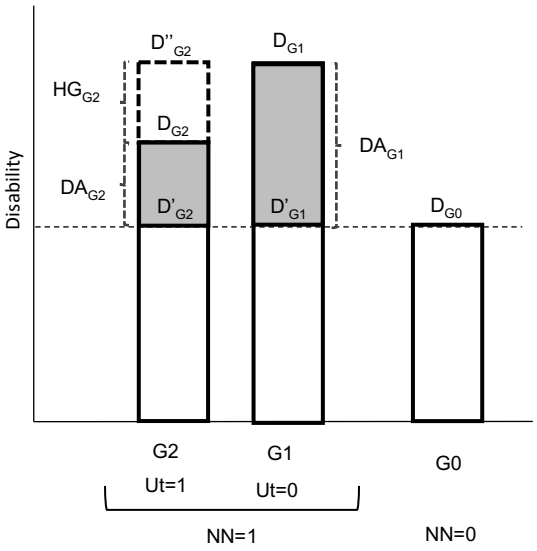
distinction between ‘burden of disability attributable’ and just ‘disability attributable’ is that the former includes the entire population, while the latter considers only people with the normative need (see below and chapter three).

The formulation of this proposal will be explained step by step, while additional concepts that complement the final approach for measuring effective coverage will also be defined. To facilitate this explanation, in Figure 4.1 I present a schematic representation of the parameters used. Additionally, Table 4.1. includes a summary of the relevant concepts and their equations.

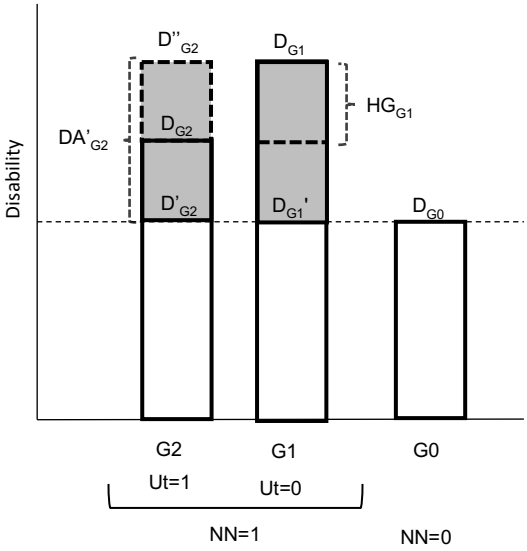
To start, Figure 4.1 Panel A depicts the expected level of disability (in a scale from 0 to 100, meaning full health and extreme disability, respectively), for a population that is split into three groups: (G0) people without the normative need, (G1) people with the normative need but who do not utilise the healthcare intervention, and (G2) people with the normative need who utilise the healthcare intervention. This representation is analogous to the figure presented in chapter three (see Figure 3.1, Panel A). The disability attributable to the normative need is represented in grey. The dotted bar depicts the disability avoided by the healthcare intervention in G2, which is equivalent to the concept of health gain. In this figure, we are assuming that people from G0, G1 and G2 are interchangeable in terms of the factors determining their level of disability, except for the normative need and the utilisation of the healthcare intervention.

Figure 4.1. Schematic representation of different parameters used to estimate effective coverage, according to the current scenario and two counterfactual ones.

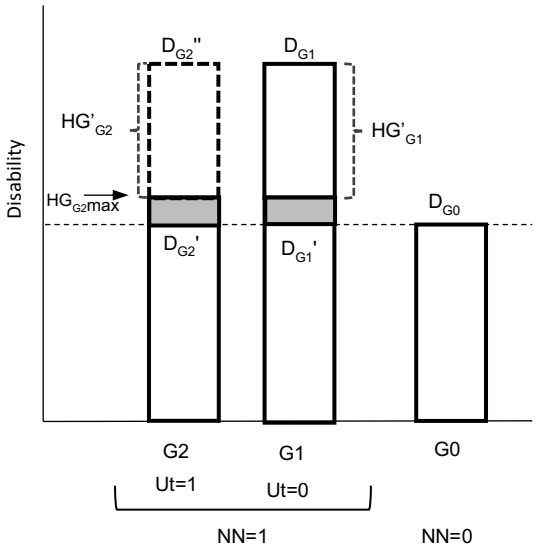
A. Actual scenario of disability



B. 'Worst' counterfactual scenario of



C. 'Adjusted by the maximum health gain' counterfactual scenario of disability



Shadows represent the amount of disability without coverage from the healthcare intervention.
NN: normative need/ Ut: utilisation/ G0: people without NN/ G1: people with NN without Ut / G2: people with NN with Ut.
D_{G0}: disability in G0/ D_{G1}: disability in G1 / D_{G1}': disability in G1 assuming they have not NN/ D_{G1}'': disability in G1 assuming they have Ut / DA_{G1}: disability attributable to NN in G1/ HG_{G2}: health gain in G2/ HG_{G2}max: maximum health gain in G2
D_{G2}: disability in G2/ D_{G2}': disability in G2 assuming they have not NN/ D_{G2}'': disability in G1 assuming they have not Ut/ HG_{G1}: health gain in G1/ HG_{G1}max: maximum health gain in G1

The first step is to calculate the fraction of disability attributable to a normative need, which is avoided by the healthcare intervention in those who use such an intervention. Similar to what was shown in chapter three for calculating the burden of disability attributable to diseases, this fraction can be expressed following a population-average or an individual-level approach. According to the population-average approach, this fraction can be written:

$$\text{Relative benefit} = \frac{HG_{G2}}{DA'_{G2}} \qquad \text{(Eq.4.1)}$$

where HG_{G2} is the health gain calculated in people from G2 (with normative need and utilisation), and DA'_{G2} is the predicted disability attributable to the normative need, also in G2, assuming they are not receiving the healthcare intervention. Note that HG_{G2} is equivalent to $DA'_{G2} - DA_{G2}$, where DA_{G2} is the current expected disability attributable to the normative need, or, in other words, the difference between the predicted level of disability in , assuming that they are not receiving the healthcare intervention, and the current expected level of disability, which is $D''_{G2} - D_{G2}$.

The health gain can also be described as the benefit of a healthcare intervention in absolute terms. In this case, because I am calculating the weight of the health gain on the total disability attributable to a normative need (i.e. a disease), it seems more convenient to call it the 'relative benefit' of the healthcare intervention. This expression is equivalent to the concept of 'exposed attributable fraction' introduced in chapter three (see Eq.3.2), or, more appropriately, the 'preventable fraction'[6, 7] since we observe a reduction in the rate of disability. However, it is important to note that this preventable fraction is within another exposed attributable fraction, which is the disability attributable to normative need.

On the other hand, following an individual-level approach, the expression can be written as:

$$\text{Relative benefit} = \frac{\sum_{k=1}^{m} (HG_{G2k}\ W_k)}{\sum_{k=1}^{m} (DA'_{G2k}\ W_k)} \qquad \text{(Eq.4.2)}$$

where the individual *k* belongs to the *m* total individuals from G2, and *W* is the sample weight from the survey design. Following chapter three, DA'_{G2} for the individual *i* is $D''_{G2i} - D'_{G2i}$ (see Figure 4.1, Panel A).

The relative benefit provides us with an idea about the effectiveness of the health intervention at the population level, in terms of the disability, level of health, or health-state utilities according to the outcome we decide to use.

The second step is to calculate a relative measure of the health coverage's impact on the total disability attributable to disease, including people with and without the healthcare intervention. For this I propose the term 'relative effective coverage' (r-EC), defined as the fraction of the disability attributable to a normative need, which is avoided given the current level of healthcare intervention use among people with the normative need.

Conceptually, the r-EC is similar to the relative benefit, but weighted by the level of utilization (i.e. the fraction of people with a normative need who utilise the healthcare intervention). In the r-EC, the numerator of the fraction corresponds to the health gain, while the denominator is the disability attributable to a normative need. However, the current measure of disability attributable to the normative need does not include the avoided disability given the healthcare intervention, which it is, therefore, necessary to add.

According to the population-average approach, r-EC can be defined as:

$$\text{r-EC} = \frac{HG_{G2}\ Ut}{DA_{G1}(1\text{-}Ut)\ +\ DA'_{G2}\ Ut} \qquad \text{(Eq.4.3)}$$

where Ut is the utilisation of the healthcare intervention, and DA_{G1} is the expected disability attributable to the normative need in G1 (i.e. people with a normative need and without utilisation).

The expression following an individual level approach is:

$$\text{r-EC} = \frac{\sum_{k=1}^{m} (HG_{G2k}\ W_k)}{\sum_{i=1}^{n} (DA_{G1i}\ W_i) + \sum_{k=1}^{m} (DA'_{G2k}\ W_k)} \qquad \text{(Eq.4.4)}$$

where the individual *i* belongs to the *n* total individuals from G1, and DA_{G1} for the individual *i* is equivalent to $D_{G1i} – D_{G1i}$' (see Figure 4.1, panel A).

The indicator of r-EC can also be understood as the contrast between the current population disability attributable to a normative need and the worst counterfactual scenario, where nobody receives the health care intervention, expressed as a fraction of the latter (see Figure 4.1, Panel B).

The r-EC gives us a relative expression about the weight of the benefit associated with the current level of coverage for a healthcare intervention, over the disability associated with a normative need. However, this does not allow us to compare the impacts of different interventions. For example, we could obtain an r-EC of 15.0% for interventions for depression, and 15.0% for interventions for diabetes, but these numbers would still not tell us which healthcare intervention was more impactful at the level of the entire population.

To cope with this problem, I propose a third step: measuring the fraction of the burden of disability attributable to a normative need in the entire population, which has been avoided through the current use of a healthcare intervention by the people with a normative need. The fraction could be called 'absolute effective coverage' (a-EC). This expression is very similar to the r-EC, but the denominator is no longer restricted to people with a normative need, but expanded to include the whole population. In other words, the a-EC corresponds to the relative benefit weighted by utilisation of the healthcare intervention and the prevalence of the normative need. And, since we are using a common denominator for different normative needs, the comparison between different healthcare interventions is more straightforward.

Following a population-average level approach, we can formulate a-EC as:

$$\text{a-EC} = \frac{HG_{G2} * NN * Ut}{NN\,[D_{G1}\,(1\text{-}Ut) + D''_{G2}Ut] + [D_{G0}\,(1\text{-}NN)]} \qquad \text{(Eq.4.5)}$$

where NN is the fraction of people with a normative need, which can be assimilated to the prevalence of a disease, D_{G0} is the expected level of disability in G0, D_{G1} is the expected level of disability in G1, and D''_{G2} is the predicted level of disability in people from G2, assuming that they are not receiving the healthcare intervention (see Figure 4.1, Panel A).

Using an individual-level approach, a-EC can be expressed as:

$$\text{a-EC} = \frac{\sum_{k=1}^{m} (HG_{G2k}\, W_k)}{\sum_{h=1}^{p} (D_{G0h}\, W_h) + \sum_{i=1}^{n} (D_{G1i}\, W_i) + \sum_{k=1}^{m} (D''_{G2k}\, W_k)} \qquad \text{(Eq.4.6)}$$

where the individual *h* belongs to the *p* total individuals from G0 (without normative need), D_{G0h} is the expected disability for the individual *h*, D_{G1i} is the expected disability for the individual *i*, and D''_{G2k} is the predicted disability for the individual *k*, assuming that the individual is receiving the healthcare intervention (see Figure 4.1, Panel A).

It is important to notice that neither r-EC nor a-EC include an expression for healthcare intervention quality. Both expressions implicitly assume that the healthcare intervention can resolve all the disability attributable to a normative need. In other words, to achieve 100% of r-EC, 100% of people with a normative need should be using the healthcare intervention (i.e. 100% coverage) and the intervention should be able to avoid 100% of the disability attributable to the normative need (100% effectiveness).

However, in some cases, current health technology cannot offer complete relief for a disease [8]. For instance, in the case of depression, almost a third of patients receiving treatment will not obtain a full relief of their symptoms [9, 10].

Moreover, according to Shengelia et al., to calculate effective coverage we need to know the 'maximum possible health gain' that an individual can expect to obtain, and not – as we have calculated up until now – the total health loss attributable to a normative need.

Consequently, we need to find a maximum health gain to use as a point of comparison for assessing the health gains associated with healthcare interventions. One option for obtaining such a maximum is to extract it from the scientific literature, first taking the precaution of translating the parameters to the scale of the outcome we are using to calculate effective coverage. Another option, as I will show in the next section, is trying to extract a maximum health gain from the distribution of predicted values of the health gain in G2 (i.e. people with a normative need who utilise the healthcare intervention), employing the same data with which we calculated the burden of disability attributable to a normative need.

Once the maximum health gain has been obtained, we can proceed to the next step, which is to calculate an indicator for the quality of the healthcare intervention. The reasoning here is similar to that of the relative benefit, but instead of using as a denominator all the disability attributable to a normative need in G2, we will use only the maximum heath gain. Consequently, in this context, 'quality' is defined as the fraction of the avoidable disability (i.e. the maximum health gain) from a normative need, which is avoided by the healthcare intervention in those who utilise the intervention.

Following the approach from Shengelia et al., we assume that any difference between the maximum health gain and the current health gain is the result of the suboptimal quality of the intervention. We have to remember that, as mentioned in chapter one, within this particular concept of quality the authors include those factors associated with implementing the healthcare intervention and the adherence to such an intervention.

A formal expression for quality, following a population-average approach, is:

$$\text{Quality} = \frac{HG_{G2}}{HGmax_{G2}} \qquad \text{(Eq.4.7)}$$

where $HGmax_{G2}$ is the maximum health gain calculated for G2. The corresponding expression for an individual-level approach is:

$$\text{Quality} = \frac{\sum_{k=1}^{m} (HG_{G2k}\ W_k)}{\sum_{k=1}^{m} (HG'_{G2k}\ W_k)} \qquad \text{(Eq.4.8)}$$

where HG'_{G2} is the vector of HG, with values at least equal to the $HGmax_{G2}$.

Figure 4.1, Panel C, depicts the counterfactual scenario 'adjusted by maximum health gain', where we are assuming that people in G1 and G2 are receiving the maximum health gain. There, the maximum health gain represents the avoidable disability given current health technology development in optimal conditions of healthcare provision.

Once we have the quality parameter, our final step is to calculate the effective coverage according to Shengelia et al.'s conceptual framework. As mentioned previously, I define effective coverage as the fraction of avoidable disability (i.e. the maximum potential health gain) attributable to a normative need (i.e. disease), which is avoided given the current level of healthcare intervention use among people with the normative need. This definition is very similar to that for r-EC. However, in this case, the denominator is no longer all the disability attributable to a normative need, but only the avoidable disability, meaning the maximum potential health gain.

Consequently, a formal expression of effective coverage following the population-average approach is:

$$\text{Effective coverage} = \frac{HG_{G2} * Ut}{(1-Ut)*HGmax_{G2} + Ut*HGmax_{G2}}$$

$$\text{Effective coverage} = \frac{HG_{G2} * Ut}{HGmax_{G2}} \qquad \text{(Eq.4.9)}$$

while the individual-level approach is expressed as:

$$\text{Effective coverage} = \frac{\sum_{k=1}^{m} (HG_{G2k}\, W_k)}{\sum_{i=1}^{n} HG'_{G1i}\, W_i + \sum_{k=1}^{m} HG'_{G2k}\, W_k} \qquad \text{(Eq.4.10)}$$

where HG'_{G1} is the vector of HG in people from G1 with values at least equal to $HGmax_{G2}$.

Table 4.1 Summary of definitions and equations of effective coverage and related concepts.

Indicator	Definition	Numerator	Denominator	Numerator	Denominator
Relative benefit	The fraction of the disability attributable to a normative need, that is avoided by the healthcare intervention, **in people with the normative need and utilisation.**	The disability attributable to a normative need that is avoided by the utilization of a healthcare intervention (i.e. the health gain).	The disability attributable to a normative need assuming the absence of the healthcare intervention, **in people with the normative need and utilisation.**	Population–average approach: HG_{G2}	Population average level approach: DA'_{G2}
				Individual–level approach: $\sum_{k=1}^{m}(HG_{G2k}\ W_k)$	Individual level approach: $\sum_{k=1}^{m}(DA'_{G2k}\ W_k)$
Relative effective coverage (r-EC)	The fraction of the disability attributable to a normative need, that is avoided given the current utilisation of a healthcare intervention, **in people with the normative need.**	The disability attributable to a normative need that is avoided by the utilisation of a healthcare intervention (i.e. the health gain).	The disability attributable to a normative need assuming the absence of the healthcare intervention, **in people with the normative need.**	Population–average level approach: $Ut*HG_{G2}$	Population average level approach: $(1\text{-}Ut)*DA_{G1} + Ut*DA_{G2}'$
				Individua –level approach: $\sum_{k=1}^{m}(HG_{G2k}\ W_k)$	Individual level approach: $\sum_{i=1}^{n}(DA_{G1i}\ W_i) + \sum_{k=1}^{m}(DA'_{G2k}\ W_k)$
Absolute effective coverage (a-EC)	The fraction of the burden of disability attributable to a normative need, that is avoided given the current utilisation of a healthcare intervention, **among people from the entire population.**	The burden of disability attributable to a normative need that is avoided by the utilisation of a healthcare intervention (i.e. the health gain).	The burden of disability in people from the entire population, assuming the absence of the healthcare intervention, **among people from the entire population.**	Population–average approach: $Ut*NN*HG_{G2}$	Population average level approach: $NN\ [D_{G1}\ (1\text{-}Ut) + D''_{G2}Ut] + [D_{G0}\ (1\text{-}NN)]$
				Individua –level approach: $\sum_{k=1}^{m}(HG_{G2k}\ W_k)$	Individual level approach: $\sum_{h=1}^{p}(D_{G0h}\ W_h) + \sum_{i=1}^{n}(D_{G1i}\ W_i) + \sum_{k=1}^{m}(D''_{G2k}\ W_k)$

Table 4.1 Summary of definitions and equations of effective coverage and related concepts. (Continuing from previous page)

Indicator	Definition	Numerator	Denominator	Numerator	Denominator
Quality	The fraction of the **avoidable** disability attributable to a normative need (i.e. the maximum health gain), that is avoided by the healthcare intervention, **among people with the normative need and utilisation.**	The disability attributable to a normative need that is avoided by the utilisation of a healthcare intervention (i.e. the health gain).	The **avoidable** disability attributable to a normative need (i.e. the maximum health gain), **among people with the normative need and utilisation.**	Population–average approach: HG_{G2}	Population–average approach: $HGmax_{G2}$
				Individual–level approach: $\sum_{k=1}^{m}(HG_{G2k}\ W_k)$	Individual–level approach: $\sum_{k=1}^{m} HG'_{G2k}\ W_k$
Effective coverage	The fraction of the **avoidable** disability attributable to a normative need (i.e. the maximum health gain), that is avoided given the current utilisation of a healthcare intervention, **among people with the normative need.**	The disability attributable to a normative need that is avoided by the utilisation of a healthcare intervention (i.e. the health gain).	The **avoidable** disability attributable to a normative need (i.e., the maximum health gain), **among people with the normative need.**	Population–average approach: $Ut*HG_{G2}$	Population–average approach: $HGmax_{G2}$
				Individual–level approach: $\sum_{k=1}^{m}(HG_{G2k}\ W_k)$	Individual–level approach: $\sum_{i=1}^{n} HG'_{G1i}\ W_i + \sum_{k=1}^{m} HG'_{G2k}\ W_k$

NN: normative need/ Ut: utilisation/ W: sample weights of the survey (the inverse of the probability to be chosen).
G0: people without NN/ G1: people with NN without Ut / G2: people with NN with Ut.
D_{G0}: disability in G0/ D_{G1}: disability in G1 / D_{G1}': disability in G1 assuming they have not NN/ D_{G1}'': disability in G1 assuming they have Ut / DA_{G1}: disability attributable to NN in G1/ HG_{G2}: health gain in G2/ $HGmax_{G2}$: maximum health gain in G2
D_{G2}: disability in G2/ D_{G2}': disability in G2 assuming they have not NN/ D_{G2}'': disability in G1 assuming they have not Ut/ HG_{G1}: health gain in G1/ $HGmax_{G1}$: maximum health gain in G1

In this section, I have defined five concepts: relative benefit, r-EC, a-EC, quality, and effective coverage, using two approaches of estimation: the population–average and the individual-level approach.

Before assessing these definitions using real data, it is necessary to discuss how to produce the estimates practically.

4.3. How to calculate effective coverage, including health gains

There are several alternatives for estimating the parameters required to calculate effective coverage. In this section, I will show which of those alternatives would be the most convenient to implement. The strengths and limitations of the selected alternatives will then be discussed later in this chapter.

Figure 4.2 illustrates the process for estimating effective coverage following the population-average approach, summarised in five steps. This process leads to the calculation of equations Eq.4.1, Eq.4.2 and Eq.4.5, which have been introduced above.

In the first step, the prevalence of the normative need and the utilisation of the healthcare intervention among prevalent cases is calculated. Then, using the estimation and its standard errors, assuming a Beta distribution, ten thousand simulated values are generated for the expected values of prevalence and utilisation. These vectors are saved for later analysis (see Figure 4.2).

Secondly, a linear regression model is formulated, including all our assumptions about the presence of a normative need, the utilisation of a healthcare intervention, and the outcome – in this case, disability. The linear regression model will be used to calculate expected values of disability and the predicted level of disability corresponding to counterfactual scenarios.

One option for the regression model's parametrisation is to create a new variable that allows us to classify survey respondents into one of three exclusive groups: G0, G1 or G2. This can also be achieved using two dummy variables, which we can call G1 and G2. A combination of G1=0 and G2=0 represents people from the group G0, i.e. without normative need. Meanwhile, G1=1 and G2=0 represent the people within the group G1, with normative need but without utilisation. Finally, G1=0 and G2=1 represents people from the group with normative need and utilisation, G2.

Another option for parametrisation is to create two non-exclusive dummy variables: one for the normative need and another for the utilisation. However, this option could make more complex the interpretation of the regression coefficients, the selection of potential confounders, and the inclusion of interaction terms.

Following the first option, the model for the expected value of disability can be expressed as follows:

$$E(Disability) = B_0 + B_1Age + B_2Sex + B_3G1 + B_4G2$$

where disability is assumed as a continuous variable. According to this model, someone with a normative need and utilisation (i.e. variable G1=0 and variable G2=1) would have an expected level of disability equal to:

$$E(Disability) = B_0 + B_1Age + B_2Sex + B_4G2$$

while someone with a normative need and without utilisation (i.e. variable G1=1, and variable G2=0) would have an expected level of disability equal to:

$$E(Disability) = B_0 + B_1Age + B_2Sex + B_3G1$$

and somebody without a normative need (i.e. variable G1=0 and variable G2=0) would have an expected level of disability equal to:

$$E(Disability) = B_0 + B_1Age + B_2Sex$$

This approach is similar to the one used to calculate the burden of disease attributable to disability presented in chapter three. It is flexible, because it allows the model to be adjusted for several confounders and/or interaction terms, whenever necessary. However, this case is a little more complex, because we are interested in isolating the effect of belonging to two different groups (i.e. G1 and G2), compared to those without a normative need (i.e. people from G0).

For example, we might be interested in adjusting the model by comorbidities. This can be justified by the fact that part of the disability present in people with a normative need (for instance, depression) could be confounded by the co-presence of a second disease (e.g. a musculoskeletal disorder). However, another justification is that having a comorbidity could be associated with a higher utilisation (or access) to the healthcare intervention addressing

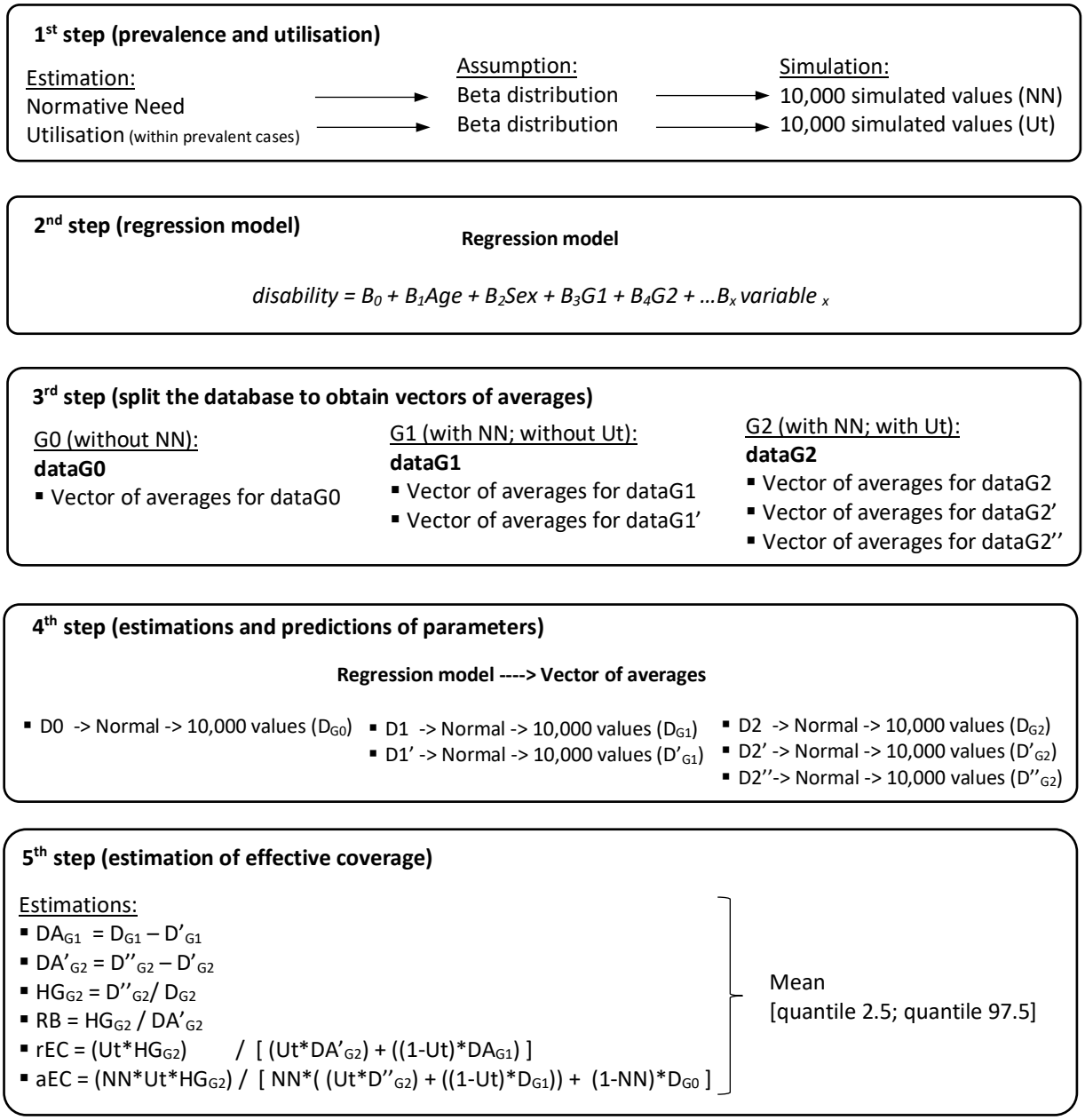
the normative need (i.e. people with depression also having a musculoskeletal disorder are more likely to access healthcare interventions for depression).

By this point, we have defined the linear regression model and continued with the process for estimating effective coverage. Then, in a third step, the database is split into three according to the groups previously defined (i.e., G0, G1, and G2), and the average value for each covariate included in the linear regression model is calculated for every group. In this way, we are generating vectors of averages that will later be used to calculate expected and predicted disability values. This step also produces the inputs for the counterfactual scenarios, changing the values of the G1 and G2 variables to 0 or 1 accordingly. The output of this stage is six vectors with average values (see Figure 4.2).

In a fourth step, the linear regression model is applied to each vector of averages, which results in the estimation of the parameters D_{G0}, D_{G1} and D_{G2} and the predicted values for D'_{G1}, D'_{G2} and D''_{G2} (see Figure 4.1, Panel A). Then, using the standard errors of these estimations and predictions, assuming a normal distribution, 10,000 simulated values are generated, which are saved for the next stage.

Finally, in the fifth step, the vectors of simulated values for normative need, utilisation (both saved from the first step), and D_{G0}, D_{G1}, D_{G2}, D'_{G1}, D'_{G2}, and D''_{G2} are used to calculate the health gain, the relative benefit, the r-EC and a-EC. Since the results are also vectors, the mean is used as a summary measure and quantiles 2.5 and 97.5 are used as uncertainty intervals.

Figure 4.2. Diagram of the implementation of effective coverage estimation using a population – average level approach.



NN: normative need/ Ut: utilisation/ DA: disability attributable/ HG: health gain/ RB: relative benefit/ rEC: relative effective coverage/ aEC: absolute effective coverage.

See the text for an explanation.

It can be seen that, following the population-average approach, the maximum health gain was not calculated. Therefore, the effective coverage was not calculated either, according to

Shingelia et al.'s framework. A proposal to estimate the maximum health gain is included in the implementation of the individual-level approach, which is presented schematically in Figure 4.3.

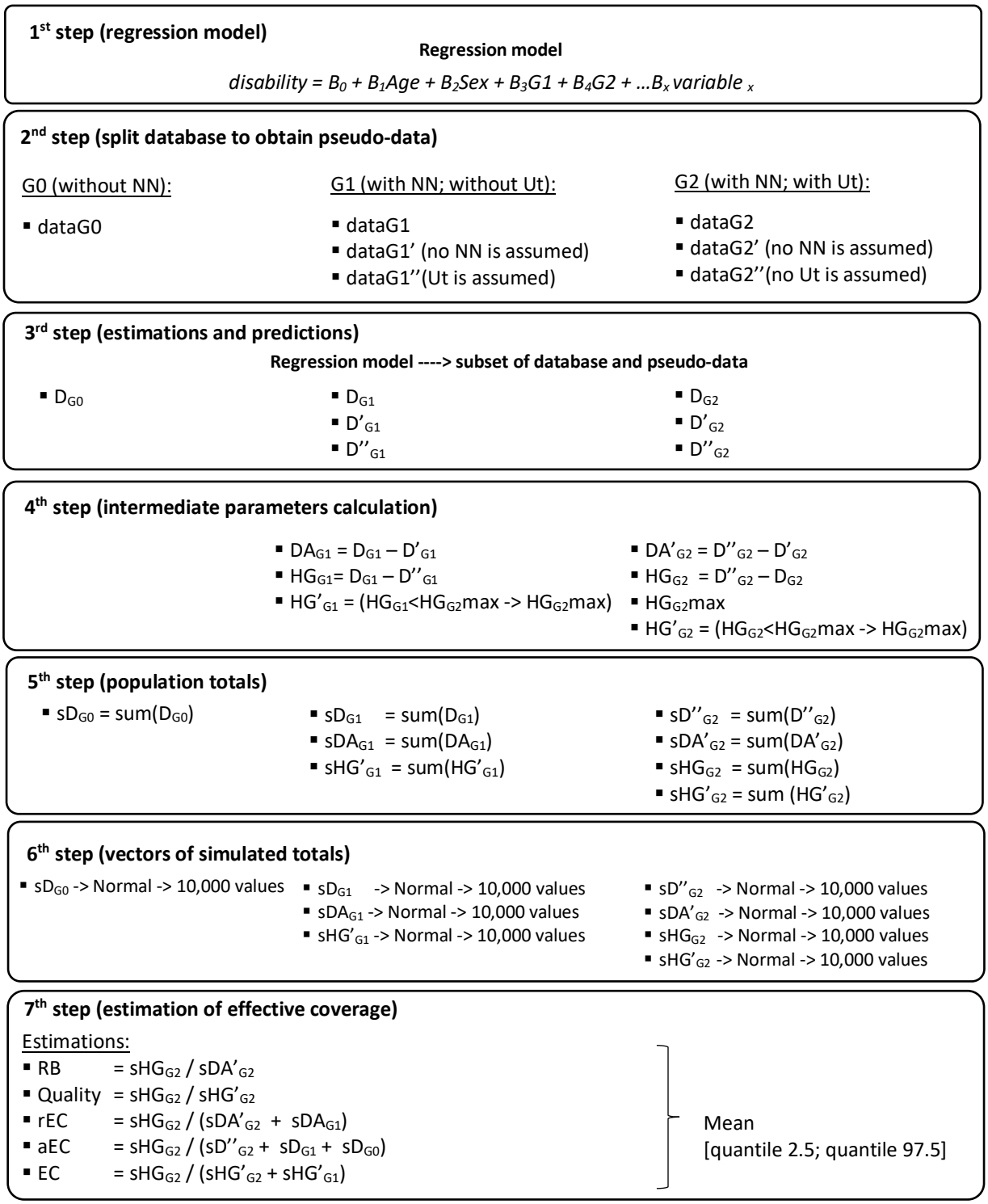
In this case, and in a similar way to the previous approach, the first step is to define the linear regression model that will include all our assumptions about the relationship between the outcome, the normative need, and the utilisation.

The second step consists of splitting the database according to the groups previously defined (i.e., G0, G1 and G2), and calling the resultant subsets 'dataG0', 'dataG1' and 'dataG2'. Additionally, four sets of pseudo-data have been generated, where the values of the dummy variables G1 and G2 are modified depending on the counterfactual scenario. Compared to the population-average approach, we are no longer modifying a vector of averages, but rather variables, into a database.

The linear regression model is applied to each set of data or pseudo-data in a third step, producing individual-level estimates and predictions for D_{G0}, D_{G1}, D'_{G1}, D''_{G1}, D_{G2}, D'_{G2}, and D''_{G2} (see Figure 4.1, Panel A).

For the fourth step, the required parameters were calculated at the individual level, such as the disability attributable to the normative need in people from G1 (i.e. DA_{G1}) and in G2, assuming they are not utilising the healthcare intervention (i.e. DA'_{G2}). The health gain and the potential health gain in G2 and G1 are also calculated respectively (i.e., HG_{G2}, HG_{G1}), as summarised in Figure 4.3.

Figure 4.3. Diagram of the implementation of effective coverage estimates, using an individual–level approach.



DA: disability attributable/ HG: health gain/ RB: relative benefit/ rEC: relative effective coverage/ aEC: absolute effective coverage/ EC: effective coverage.

See the text for an explanation.

The maximum health gain in people with a normative need and with healthcare utilisation (dataG2) is also calculated in this step. For this calculation, I propose choosing an arbitrary value of HG that can be used as a maximum value to assess the health gains obtained from each individual. For example, we can use the HG corresponding to the quantile 0.90, extracted from a set of HGs calculated for the individuals with normative need and utilisation. However, the HG at the individual level in people from G2 is calculated from the difference between the predicted disability D_{G2}'' and the expected disability D_{G2} (i.e. HG = D_{G2}'' – D_{G2}). When we use a linear regression model, a unique value for HG is obtained for all the individuals. And, when the linear regression model is applied on dataG2, that value is equivalent to the linear regression coefficient for the variable G2.

To resolve this issue, I propose including in the linear regression model the interaction terms between the variable G2 and those variables that could modify the quality of the healthcare intervention – that is the health gain – in terms of avoided disability. As mentioned above, the quality of the intervention might be affected by two factors: namely adherence and implementation. The variables for these interactions can include the geographical area, education, income, sex, age, and comorbidities, among others. In this way, we can generate specific predicted values of HG for each individual, according to his or her own set of covariates.

This approach can increase the collinearity of the variables in the regression model, enlarging standard errors. However, a stepwise process of variable selection can be implemented to reduce this problem (see the next section of this chapter).

Generating a set of HG for choosing a maximum health gain causes another inconvenience, which is the increasing complexity in the analysis of which variables should be included in the linear regression. The variables to include must be thought of as either confounders of the relationship between normative need/utilisation and disability, or as potential modifiers of the utilisation effect on disability. This aspect will be discussed later in the chapter.

After a maximum health gain in people from G2 (i.e. HG_{G2}max) has been chosen, it might be possible to calculate the potential HG that the intervention could offer if people receive such health benefits. For this, at the individual level, within the vectors HG_{G1} and HG_{G2} I replace all the values lower than the HG_{G2}max with this new value, generating two new vectors: HG_{G1}' and HG_{G2}', respectively.

The values of the vectors' HG_{G2}' and HG_{G1}' are not necessarily identical to the HG_{G2}max, because the threshold used to choose the maximum health gain can be lower than the quantile 1. Selecting a threshold lower than 1 allows a degree of variability in the HG. This assumes: 1) a normative tolerable difference in the quality of the healthcare intervention, or 2) a difference in health gains due to factors different to the quality, such as higher effectiveness of healthcare interventions in some subgroups of the population. For example, if the HG_{G2}max using a threshold equal to quantile 0.9, is 3.5 points of disability (on a scale between 0 and 100), it means that all the HGs inferior to 3.5 are sub-optimal. In other words, they come from a healthcare intervention of lower quality. On the contrary, in the few cases where the HG values are higher than 3.5, it is assumed to be a tolerable variability in the quality, or else a difference that can be attributed to factors other than quality. Therefore, using an arbitrary threshold for choosing a maximum HG makes it desirable to explore the results from different values.

Finally, it is important to notice that the proposed parameterisation of the HGs on a scale of disability promotes the calculation of values higher than zero, even though we expect the healthcare interventions to avert disability.

The fifth step in estimating effective coverage following an individual-level approach is to add all the scores from the different parameters to estimate the population totals (see Figure 4.3).

Then, in the sixth step, once the totals are estimated using their means and the standard errors, ten thousand simulated values are generated, assuming a normal distribution for each parameter. The output of this process is eight vectors that will be used in the final step: sD_{G0}, sD_{G1}, sDA_{G1}, sHG_{G1}', sD_{G2}'', sDA_{G2}', sHG_{G2}, and sHG_{G2}'.

The seventh and final step is to calculate the outcomes, such as the relative benefit, the quality, r-EC, a-EC and effective coverage. Since the results correspond to vectors, the means and quantiles 2.5 and 97.5 are extracted.

The above implementation is consistent with the formulations Eq.4.2, Eq.4.4, Eq.4.6, Eq.4.8, and Eq.4.10.

On the other hand, the value for $HG_{G2}max$, calculated through the individual-level approach, could also be used in the population-average approach to indirectly estimate the quality (Eq. 4.7) and effective coverage, according to the framework of Shengelia et al. (Eq. 4.9).

4.4. Testing the procedure to calculate effective coverage including health gains

To test the concepts and procedures presented here, I will use the second Chilean National Health Survey, the same database used in chapters two and three. This survey was carried out between 2009 and 2010 (Ch-NHS 2009-2010), in people older than 14. The survey contains the four elements required for this analysis: 1) items or questions to identify diseases; 2) information to assess the access to healthcare; 3) questions to measure the level of health/disability in individuals, and 4) several variables that can be used to adjust the relationship between health/disability, the presence of a disease and the access to healthcare. A detailed methodology of Ch-NHS 2009-2010 is available in chapter two.

4.4.1. Disability outcome

The psychometric properties of the scale that measures the disability of individuals were described in Chapter two. In that chapter, the methods used to calculate the latent variable of disability were also shown. The variable is anchored in extreme values of 0, meaning full health, and 100, meaning total disability; and inquiries about the disability concerned the previous 30 days.

4.4.2. Normative need and comorbidities

In the third chapter, five diseases were used as examples to demonstrate the process of estimating the attributable burden of disability. Among them, only one normative need – depression – was chosen to test the proposal to estimate effective coverage. The remaining diseases – hypertension, diabetes, chronic musculoskeletal pain, and chronic respiratory symptoms – were left as covariables to include in the models.

The analysis of only one normative need will allow us to explore the procedure's assumptions in more detail and provide a better understanding of the results. In Chapter Three, the depressive episode showed a high prevalence and the greatest burden of attributable disability.

Cases with a depressive episode in the previous 12 months were detected using the questionnaire CIDI-SF and the DMS-IV criteria [11-13] (see chapter two for details). The number of depressive symptoms, ranging from none to eight, was used to measure the severity of the disease.

The operational definition for other diseases was described in chapters two and three.

4.4.3. Utilisation

The Ch-NHS 2009-2010 measured the utilisation of healthcare interventions for depression with two items: ’Have you ever been treated for depression?’ (i.e. ever treated), and ‘Have you been taking any medication in the past two weeks for depression?’ (i.e. medication in the last two weeks). Both items in the Ch-NHS questionnaire are placed after this question: ‘Has a doctor or physician ever told you that you have or suffer from depression?’.

The usage items are rather general, especially the first, which could include any kind of intervention for depression. However, we can assume that the previous question will frame the answer in a certain way, establishing that the diagnosis was made by a physician.

On the other hand, the normative need is measured for the previous 12 months. Therefore, the 'ever treated’ item could bias the estimates, since it can include a positive answer from those who received treatment for a previous episode, but not the current one. In that case, the health gain attributable to the healthcare intervention can be underestimated, potentially biasing all the indicators of effective coverage.

The item about ‘medication in the last two weeks’ is no better, since it restricts the healthcare interventions to medication alone. Since depressive episodes can be treated with interventions other than medication – especially milder cases [14] – the utilization can again be underestimated. Moreover, according to the natural history of a depressive episode, remission is expected within a year [9], meaning that again we might underestimate the level of utilisation and the health gains.

For this analysis, I chose ‘ever treated’ as a proxy for utilisation as the main scenario. Since the Ch-NHS also asks about the date of the first diagnosis for an episode of depression, those

cases with a first diagnosis more than two years ago were excluded from the analysis for the second scenario. The usage defined as 'treatment during the last two weeks' was employed in a third analytical scenario. Finally, in a fourth scenario, the same definition as the third scenario was used, but those non-severe cases (i.e. with less than seven symptoms out of eight) were excluded.

4.4.4. Other covariates

Other covariates included in the analysis were age, sex, level of education (three categories: < 8 years, 8 – 12 years, > 12 years of education), marital status (four categories: married or cohabiting, divorced or separated, widowed, single), and the quintile of household income per capita.

4.4.5. Regression models

In the previous section, I showed that the regression model specification is an essential stage in the process of calculating the effective coverage of healthcare interventions. In this case, a linear regression model, assuming disability as a continuous variable, was used.

According to the procedure described previously, G1 and G2's dummy variables were created and included in the linear regression model. G1 adopted the value one in those cases with a depressive episode but without treatment; otherwise the value was zero. Meanwhile, G2 took the value one in those cases of a treated depressive episode. Cases without a depressive episode during the last 12 months (i.e. G1=0 and G2=0), were used as the category of reference.

The calculation of effective coverage was tested through two linear regression models. The first included G1 and G2 and potential confounders (i.e. model 1), while the second included interactions terms as well (i.e. model 2).

Age, sex, marital status, level of education, quintile of income, and comorbidities were included in the model as potential confounders for the relationship between disability and depression, or between disability and the use of healthcare interventions [15-17]. The

disability associated with the access and the potential health gain (i.e. quality) from a healthcare intervention can be different according to some of these factors [18, 19]. As was mentioned previously, interactions terms were also included between G1, G2 and other sociodemographic variables and comorbidities. A final model was selected after a stepwise procedure.

As a variable to be included in the linear regression models, the severity of the normative need requires special consideration. People who seek access to a healthcare intervention can have a higher level of initial disability than people who do not, and that level can be correlated with the level of disability after treatment. That higher level of disability may be explained by age, sex, socioeconomic status, and comorbidities; factors that can add to disability, independent of the presence of a normative need. However, it can also be due to factors inherent to the normative need, such as the severity of the health condition [20]. Additionally, the variable of severity can also interact with the variable G2, which marks disability associated with access to treatment [21, 22]. The number of symptoms, as mentioned above, was used as a measure of the severity of the depressive disorder. The CIDI-SF asks about the presence of symptoms at any moment during the last 12 months, and I am assuming that a higher number of these symptoms occurred before the access to treatment.

When a marker of the severity of the disease is included in the linear regression model, some caution is required for calculating DA_{G1} and DA'_{G2}, either in the population-average or individual-level approaches, especially when the vector of averages and the pseudo-data are generated. In both cases, it is necessary to create the counterfactual of people with a depressive episode, assuming that these individuals do not have this normative need. In order to do that, the average number of symptoms observed in people without the disease was used.

4.4.6. Statistical procedures

Results following the population-average and the individual-level approaches were obtained using linear regression models 1 and 2. In the linear regression model 2, full two-way interactions between G1 and covariables, and G2 and covariables, were implemented. This included quadratic and cubic terms for the variable age. Then, the number of variables was reduced through a backward selection procedure [23].

The population-average level and individual level approaches were implemented through functions developed for the statistical software R. Using such functions greatly facilitated the extraction of intermediate results. The syntax for the main and secondary functions is available in the supplementary material (see Main functions for R S4.1).

The sample weights from the design of the survey were considered in all the analyses. The procedures were carried out using the statistical software R 3.5.0 and its package *survey*.

4.5. Results

From the 5,412 people surveyed, 4,274 had complete data for all variables included in the analysis. Missing data was mainly for glycemia and blood pressure, which was required to identify diabetes and hypertension cases, respectively. The frequency of missing data for each variable is presented in the supplementary material, Table S4.1. No imputation on the missing data was implemented, because I am interested in testing concepts and procedures rather than extrapolating for the entire population.

The observed prevalence of a depressive episode in the previous 12 months (i.e. normative need) was 18.1% [16.1 – 20.1]. From people identified with depression, 49.6% [43.6 – 55.6] reported having ever received treatment (i.e. utilisation). The description of the whole sample is presented in Table 4.2, with separate subgroups for people without depression (i.e. G0), people with depression and without treatment (i.e. G1), and with depression and with treatment (i.e. G2).

The disability score was higher in the last two subgroups. People from G1 and G2 tended to be concentrated in the middle age range. Among people who had experienced a depressive episode, there was a greater number of women, especially among those receiving treatment. In addition, there was a trend for people with less education and lower incomes, both with and without treatment. The presence of comorbidities, such as chronic musculoskeletal symptoms and chronic respiratory symptoms, was more common in people with a depressive episode than in people without. People with depression and treatment showed a higher number of symptoms than people without treatment.

The results from regression models 1 and 2 for disability are presented in Table 4.3, including the regression coefficients from the bivariate analysis. Age and sex are associated with higher levels of disability. Differences in disability by marital status disappeared after controlling for confounders in model 1. People in the lowest category of education showed a higher level of disability, while those in the higher income quintiles were associated with lower disability levels. All comorbidities showed positive and significant coefficients. The severity of the depressive episode, measured as the number of symptoms, showed a direct association with the level of the disability. According to Model 1, each additional symptom was associated with an increase of 1.9 [1.3 – 2.6] points in the total disability score.

In the bivariate analysis, G1 and G2 were strongly associated with a higher level of disability than G0. People with depression but without treatment showed 13.1 [10.5 – 15.6] additional points of disability compared with people from G0, while people with depression and with treatment scored 14.5 [12.0 – 17.0] points higher for disability (see Table 4.3). In Model 1, after adjustment by potential confounders, the magnitude of the regression coefficients for G1 and G2 diminished towards statistical insignificance (i.e. the confidence intervals overpass the zero). However, that is explained because the variability of the depression's severity captured the association between disability and G1 and G2. If the severity of depression in Model 1 had not been adjusted, the regression coefficient for G1 and G2 would have been 11.9 [9.8 – 14.1] and 11.7 [9.5 – 14.0], respectively (data not shown in Table 4.3).

Table 4.2. Description of the sample, according to three subgroups of the population: without a depressive episode in the previous 12 months (G0); with a depressive episode, but without treatment (G1); with a depressive episode and with treatment (G2). Chilean National Health Survey 2009-2020 (n=4,447).

	Whole sample (n= 4,447)			G0 (without depression); n=3.712		G1 (with depression. without treatment); n=381		G2 (with depression. with treatment); n=354	
	n	**%. or mean**	**CI**	**% or mean**	**CI**	**% or mean**	**CI**	**% or mean**	**CI**
Disability (mean)	-	32.1	[31.4 - 32.9]	29.4	[28.6 - 30.1]	44.0	[41.5 - 46.5]	45.4	[42.9 - 47.8]
Age (mean)	-	41.5	[40.7 - 42.4]	41.5	[40.5 - 42.4]	40.6	[38.4 - 42.9]	43.0	[40.5 - 45.4]
Age (%)									
<35 yo	1,275	38.5	[35.9 – 41.0]	40.0	[37.1 - 42.8]	34.1	[26.3 - 41.8]	29.5	[21.8 - 37.3]
35 - 49 yo	1,183	29.9	[27.5 - 32.2]	27.8	[25.2 - 30.3]	38.0	[29.9 - 46.2]	40.6	[31.5 - 49.7]
50 - 64 yo	1,013	19.9	[18.0 - 21.9]	19.6	[17.4 - 21.8]	22.3	[15.4 - 29.2]	20.4	[14.1 - 26.7]
> 64 yo	803	11.7	[10.4 - 13.1]	12.7	[11.1 - 14.3]	5.6	[3.4 - 7.8]	9.5	[5.4 - 13.5]
Sex (% female)	2,579	52.2	[49.6 - 54.8]	47.4	[44.6 - 50.2]	62.2	[53.9 - 70.6]	86.3	[79.9 - 92.8]
Marital status (%)									
Married/ cohabiting	2,412	55.8	[53.3 - 58.4]	55.7	[52.9 - 58.5]	59.1	[51.2 - 67]	53.6	[44.8 - 62.3]
Divorced/ Separated	397	7.2	[6.0 - 8.5]	6.3	[5.0 - 7.5]	7.4	[3.4 - 11.4]	15.7	[9.1 - 22.2]
Widowed	426	5.2	[4.4 – 6.0]	5.0	[4.2 - 5.9]	4.8	[2.6 - 7.1]	7.5	[3.7 - 11.3]
Single	1,039	31.7	[29.3 - 34.2]	33.0	[30.2 - 35.8]	28.7	[21.4 - 36]	23.3	[16.0 - 30.6]
Education (%)									
> 12 years	778	22.4	[20.1 - 24.6]	23.1	[20.6 - 25.7]	14.1	[8.4 - 19.8]	23.8	[16.1 - 31.6]
8 - 12 years	2,349	58.0	[55.5 - 60.6]	57.8	[55.0 - 60.6]	62.9	[54.8 - 70.9]	55.3	[46.6 - 63.9]
< 8 years	1,147	19.6	[17.8 - 21.4]	19.1	[17.1 - 21]	23.0	[15.9 - 30.1]	20.9	[14.7 – 27.0]
Quintile of Income pc									
(poorer) Quintile 1	730	19.9	[17.7 – 22.0]	18.7	[16.4 - 21]	25.1	[18.1 - 32.1]	25.1	[16.2 - 33.9]
Quintile 2	939	22.8	[20.7 - 24.9]	22.1	[19.9 - 24.4]	23.0	[15.8 - 30.2]	28.6	[21.1 - 36.2]
Quintile 3	752	18.9	[16.8 – 21.0]	19.3	[17.0 - 21.7]	19.8	[13.7 - 25.8]	14.2	[8.9 - 19.5]
Quintile 4	941	21.7	[19.6 - 23.9]	22.1	[19.7 - 24.4]	20.5	[13.3 - 27.7]	20.0	[13.2 - 26.8]
(richer) Quintile 5	912	16.7	[14.9 - 18.4]	17.8	[15.8 - 19.8]	11.6	[7.0 - 16.3]	12.1	[6.9 - 17.3]
Hypertension (%)	1,450	28.0	[25.9 - 30.2]	27.8	[25.4 - 30.1]	30.3	[22.2 - 38.4]	28.5	[21.1 - 35.9]
Diabetes (%)	354	6.9	[5.8 - 8.1]	6.3	[5.1 - 7.5]	10.4	[5.8 – 15.0]	9.1	[4.8 - 13.4]
Chronic respiratory symptoms (%)	363	9.6	[7.9 - 11.2]	8.0	[6.2 - 9.7]	18.1	[11.5 - 24.7]	15.6	[9.3 - 21.9]
Chronic musculoskeletal pain(%)	1,557	31.1	[28.8 - 33.3]	26.5	[24.3 - 28.8]	47.2	[39.0 - 55.4]	56.2	[47.6 - 64.8]
Depresive episode (%)	715	18.1	[16.1 - 20.1]	0	-	100	-	100	-
Number of depressive symptoms (mean)	-	1.3	[1.2 – 1.5]	0.2	[0.1 - 0.3]	6.3	[6.2 - 6.5]	6.7	[6.5 - 6.8]

CI: confidence intervals 95%

Table 4.3. Coefficients for disability from bivariate and multivariate linear regression models. Data from the Chilean National Health Survey (n=4,447).

	-		Model 1		Model 2	
	Bivariate		Multivariate		Multivariate	
	Coef	**CI**	**Coef**	**CI**	**Coef**	**CI**
Intercept	-	-	17.0	[14.2 to 19.8]	14.1	[5.0 to 23.1]
Age (each 10 yo)	2.9	[2.6 to 3.3]	1.8	[1.3 to 2.3]	5.5	[-0.5 to 11.5]
Sex (female)	7.2	[5.8 to 8.7]	3.9	[2.7 to 5.1]	3.8	[2.6 to 5.0]
Marital status						
Married/ cohabiting	0.0	-	0.0	-	0.0	-
Divorced/ Separated	4.4	[1.5 to 7.3]	0.1	[-2.6 to 2.9]	0.8	[-2.5 to 4.0]
Widowed	9.6	[7.2 to 12.0]	0.0	[-2.3 to 2.4]	0.0	[-2.6 to 2.7]
Single	-7.3	[-8.8 to -5.8]	-0.8	[-2.3 to 0.7]	-1.0	[-2.7 to 0.8]
Education						
> 12 years	0.0	-	0.0	-	0.0	-
8 - 12 years	-2.4	[-3.9 to -0.8]	0.3	[-1.2 to 1.9]	0.6	[-1.0 to 2.1]
< 8 years	8.8	[6.9 to 10.7]	3.6	[1.4 to 5.8]	3.3	[1.1 to 5.6]
Quintile of Income pc						
(porer) Quintile 1	0.0	-	0.0	-	0.0	-
Quintile 2	1.7	[-0.1 to 3.4]	0.4	[-1.5 to 2.4]	-0.5	[-2.5 to 1.6]
Quintile 3	1.0	[-0.9 to 2.9]	0.4	[-1.7 to 2.5]	-0.7	[-2.9 to 1.5]
Quintile 4	-1.8	[-3.6 to 0.1]	-1.1	[-3.1 to 0.9]	-1.7	[-3.7 to 0.4]
(richer) Quintile 5	-2.9	[-4.6 to -1.2]	-1.9	[-3.9 to 0.2]	-2.2	[-4.2 to -0.3]
Hypertension	7.6	[5.9 to 9.3]	1.5	[-0.1 to 3.2]	1.1	[-0.7 to 2.8]
Diabetes	10.9	[8.4 to 13.5]	3.9	[1.4 to 6.3]	4.7	[2.1 to 7.4]
Chronic respiratory symptoms	8.5	[5.8 to 11.2]	6.1	[3.7 to 8.5]	5.8	[3.3 to 8.2]
Chronic musculoskeletal pain	10.4	[9.0 to 11.8]	5.2	[4.0 to 6.4]	6.0	[4.7 to 7.4]
G0 (without depression)	0.0	-	0.0	-	0.0	-
G1 (with depression, without treatment)	13.1	[10.5 to 15.6]	0.0	[-4.3 to 4.3]	-7.1	[-14.8 to 0.5]
G2 (with depression, with treatment)	14.5	[12.0 to 17.0]	-0.8	[-5.2 to 3.6]	6.6	[-5.8 to 19.0]
Severity of depression (each symptom)	2.4	[2.1 to 2.7]	1.9	[1.3 to 2.6]	2.2	[1.6 to 2.8]
Age ^2	-	-	-	-	-1.1	[-2.4 to 0.2]
Age ^3	-	-	-	-	0.1	[0.0 to 0.2]
Age:G1	-	-	-	-	1.0	[-0.4 to 2.4]
Marital status(3):G1	-	-	-	-	-7.9	[-14.2 to -1.7]
Quintile of Income pc (2):G1	-	-	-	-	5.2	[0.8 to 9.7]
Quintile of Income pc (3):G1	-	-	-	-	5.4	[0.2 to 10.6]
Quintile of Income pc (4):G1	-	-	-	-	2.7	[-1.9 to 7.3]
Hypertension:G1	-	-	-	-	3.1	[-1.9 to 8.1]
Diabetes:G1	-	-	-	-	-3.1	[-9.0 to 2.8]
Chronic Musculoskeletal Pain:G1	-	-	-	-	-3.4	[-6.7 to 0]
Age:G2	-	-	-	-	1.3	[0.0 to 2.5]
Marital status(2):G2	-	-	-	-	-2.8	[-8.5 to 2.8]
Marital status(3):G2	-	-	-	-	-6.9	[-14.7 to 1.0]
Quintile of Income pc (3):G2	-	-	-	-	3.0	[-2.5 to 8.4]
Chronic Musculoskeletal Pain:G2	-	-	-	-	-4.3	[-8.6 to 0.1]
Severity of depression: G2	-	-	-	-	-1.8	[-3.7 to 0.1]

Coef: coefficient/ CI: confidence intervals 95%.

Overall, we can see that the difference between the coefficients for G1 and G2 is rather small. G2 tended to be higher than G1 in the bivariate analysis, but that relationship is inverted after adjustment by covariables. Results from model 2 are more difficult to interpret, due to the presence of interaction terms, but they remain consistent with those in model 1.

Table 4.4 shows the outputs of the proposal to calculate effective coverage and intermediate parameters following the population-average level approach based on models 1 and 2. In

people with depression and receiving treatment, the expected disability, according to model 1, is 45.3 [43.2 – 47.5] (i.e. D_{G2}). The predicted disability in the same population, assuming no depression, is 33.6 [32.7 – 34.6] (i.e. D'_{G2}), and assuming the absence of treatment is 46.2 [44.3 – 48.3] (i.e., D''_{G2}). Therefore, the disability attributable to depression is 12.6 [10.5 – 14.6] (i.e. $DA'_{G2} = D''_{G2} – D'_{G2}$), and the health gain calculated for the treatment is 0.8 [-2.0 to 3.6] units of disability (i.e. $HG_{G2} = D''_{G2} – D_{G2}$).

Table 4.4. Effective coverage for depressive episode and other parameters estimated using the population – average level approach based on linear regression models 1 and 2, using data from the Chilean National Health Survey 2009-2020 (n=4,447).

	Population – average approach			
	Model 1		Model 2	
	mean	**UI**	**mean**	**UI**
D_{G0}	29.4	[28.7 to 30.0]	29.1	[28.0 to 30.2]
D_{G1}	44.0	[42.1 to 45.9]	43.9	[42.2 to 45.7]
D'_{G1}	32.0	[31.2 to 32.8]	32.0	[30.8 to 33.2]
D''_{G1}	-	-	-	-
DA_{G1}	12.0	[9.9 to 14.1]	11.9	[9.9 to 14.0]
D_{G2}	45.3	[43.2 to 47.5]	45.1	[42.9 to 47.3]
D'_{G2}	33.6	[32.7 to 34.5]	33.5	[32.3 to 34.7]
D''_{G2}	46.2	[44.3 to 48.0]	45.8	[44.2 to 47.4]
DA'_{G2}	12.6	[10.5 to 14.6]	12.3	[10.2 to 14.3]
HG_{G1}	-	-	-	-
HG_{G2}	0.8	[-2.0 to 3.6]	0.7	[-2.1 to 3.5]
HG_{G2}max	-	-	-	-
Relative benefit (%)	5.9	[-18.3 to 26.7]	5.1	[-18.5 to 26.6]
Quality (%)	-	-	-	-
Relative - effective coverage (%)	3.1	[-8.8 to 14.4]	2.7	[-8.9 to 14.0]
Absolute - effective coverage (%)	0.2	[-0.6 to 1.0]	0.2	[-0.6 to 1.0]
Effective coverage (%)	-	-	-	-

UI: uncertainty intervals, representing quantiles 2.5 and 97.5 from resultant distributions

DA: disability attributable/ HG: health gain

In other words, these results suggest that, from the 12.6 units of disability attributable to depression, only 0.8 are being avoided in people who are receiving treatment. In relative terms, that is equivalent to 5.9% (i.e. the relative benefit). Since the coverage of the treatment for depression is approximately 50%, the r-EC is close to half the relative benefit

– in this case just 3.1%. When we move this value to the entire population, it represents just 0.2% [-0.6 to 1.0] of the total disability (i.e. a-EC).

There are not many differences in central estimations for the magnitude or the uncertainty, using either of the linear regression models 1 or 2.

However, the size of the uncertainty intervals is noticeable, especially for the outputs that are expressed as a fraction: relative benefit, r-EC and a-EC. However, more problematic is that certain boundaries have negative values. This is a consequence of the assumption of normal distribution applied to the health gain (i.e. HG_{G2} parameter), which is derived from the subtraction between D''_{G2} and D_{G2}, which are also assumed to be normally distributed. Permitting negative numbers in HG is equivalent to assuming that the intervention could be harmful in terms of disability. Further discussion about this assumption and its implications will be given in the next section.

Table 4.5 presents the outputs of the procedure to calculate effective coverage, following the individual-level approach applied to linear regression models 1 and 2. The table shows the mean and confidence intervals of the expected and predicted values D_{G0}, D_{G1}, D'_{G1}, D''_{G1}, DA_{G1}, D_{G2}, D'_{G2}, D''_{G2} and DA'_{G2}. The information was extracted from steps three and four, according to the diagram presented in Figure 4.3. The magnitude of the parameters is similar in models 1 and 2, and similar to those obtained using the population-average approach. The most important difference is observed in the size of the uncertainty interval. In the individual-level approach, the uncertainty tends to be larger, since the estimations were performed on subsamples of the population represented in the data.

The HG calculated for people with depression and receiving treatment was almost identical to those shown using the population-average approach. However, in the case of the individual-level approach, applied to model 1, we can see that the uncertainty is null (i.e. confidence intervals are the same as the central point estimation). This is because there is no variation in HG_{G2} among the individuals, since it was not included in the interaction terms between G2 and other covariates. HG_{G2} in model 1 is dependent on the regression coefficient for variable G2 (see Table 4.2)

Table 4.5. Effective coverage for depressive episodes and other parameters, estimated using the individual-level approach applied to linear regression models 1 and 2. Data from the Chilean National Health Survey 2009-2020 (n=4,447).

	Individual level approach					
	Model 1		Model 2		Model 2 (totals/100)	
	mean	**UI**	**mean**	**UI**	**mean**	**UI**
D_{G0}	31.2	[19.6 to 46.3]	31.3	[19.8 to 48.1]	2,438,831	[2,330,614 – 2,547,049]
D_{G1}	45.2	[32 to 59.7]	45.2	[31.3 to 61.9]	407,258	[352,879 – 461,638]
D'_{G1}	33.3	[21.4 to 46.3]	33.4	[21.5 to 47.7]	-	-
D''_{G1}	44.3	[31.2 to 58.9]	45.5	[33.5 to 61.3]	-	-
DA_{G1}	11.8	[9.4 to 15.2]	11.8	[3.1 to 19.8]	110,151	[925,39 – 127,763]
D_{G2}	46.4	[33.4 to 60.4]	46.6	[35.4 to 62.4]	-	-
D'_{G2}	34.8	[23.3 to 48.3]	34.9	[23.4 to 49.8]	-	-
D''_{G2}	47.2	[34.2 to 61.2]	47.3	[32.3 to 64.3]	419,256	[361,361 – 477,151]
DA'_{G2}	12.5	[9.4 to 15.2]	12.4	[2.7 to 21.4]	111,687	[95,775 – 127,599]
HG_{G1}	0.8	[0.8 to 0.8]	-0.3	[-5.8 to 6.2]	-	-
HG_{G2}	0.8	[0.8 to 0.8]	0.7	[-5.9 to 7.4]	6,420	[1,475 – 11,365]
HG_{G2}max	0.8	-	5.3	-	-	-
Relative benefit (%)	6.4	[5.2 to 7.8]	5.8	[1.4 to 10.4]	-	-
Quality (%)	100.5	[81.7 to 122.5]	13.1	[3.1 to 23.6]	-	-
Relative - effective coverage (%)	3.3	[2.7 to 3.8]	2.9	[0.7 to 5.2]	-	-
Absolute - effective coverage (%)	0.2	[0.2 to 0.3]	0.2	[0.0 to 0.3]	-	-
Effective coverage (%)	49.7	[41.5 to 58.6]	6.5	[1.6 to 11.6]	-	-
$HG'_{G1} + HG'_{G2}$	-	-	-	-	99,159	[89,661 – 108,837]

UI: uncertainty intervals. Some parameters represent confidence intervals, others quantiles 2.5 and 97.5 from the resultant distributions (see the text).

DA: disability attributable/ HG: health gain

Comparing the results for the relative benefit, r-EC, and a-EC, again we see similarities between models and approaches. The greater distinction is that the uncertainty intervals are considerably smaller for these results. In part, this is because the estimates of the total disability units (step five from the diagram in Figure 4.3) do not include the uncertainty from predictions at the individual level [24].

Regarding quality and effective coverage, in the case of the results from model 1, since there is no variation in HG_{G2} (i.e. 0.8 [0.8 – 0.8]], the maximum HG was fixed using the same value.

Consequently, the quality was estimated as 100% and the effective coverage became equivalent to the raw coverage estimation: 49.7% [41.5 – 58.6].

However, the individual-level approach applied to the linear regression model 2 allows a 'pseudo-variation' of the HG across individuals with depression receiving treatment. This 'pseudo-variation' is obtained thanks to the presence of the interaction terms. According to the results of the linear regression model 2 from Table 4.3, five variables interacted with the variable G2: age, marital status, quintile of income, chronic musculoskeletal pain, and the severity of the depressive episode. These interactions determined different values of HG_{G2} for each individual – values that accounted for the size of the uncertainty intervals (i.e. 0.7 [-5.9 to 7.4]).

Using as a threshold the percentile 0.9 of the HG_{G2} values, a maximum HG equivalent to 5.3 units of disability was determined. This allowed counterfactuals to be generated, where the benefit of the treatment is at least equal to that magnitude (i.e. HG'_{G1} and HG'_{G2}).

Given the new counterfactuals of maximum HG, the quality of the treatment could be calculated from the fraction between HG_{G2} and the maximum HG_{G2}: 13.1% [3.1 – 23.6]. After obtaining the maximum HG_{G2}, the estimated effective coverage was: 6.5% [1.6 – 11.6]. That is, 6.5% of the avoidable disability attributable to depression is avoided through treatment for that condition.

Additionally, it can be observed that effective coverage is equivalent to the product between the raw coverage and the quality: 49.6% * 13.1% = 6.5%. Furthermore, the maximum HG calculated by the individual-level approach can be used in the equations Eq.4.7 and Eq.4.9, presented above, to obtain an approximation of the quality and effective coverage for the population-average approach: 13.3% [-42.0 to 70.5] and 6.6% [-20.7 to 35.1], respectively. Once again, the central values (i.e. the point estimation) of quality and effective coverage are equivalent between the population-average and the individual-level approaches.

In Table 4.5, some of the parameters from the individual-level approach applied to model 2 are also shown. They are expressed as person-years of extreme disability, and assume the constant presence of depression and use of healthcare interventions in a whole year (see discussion in chapter three). The estimated health gain produced by the treatment of

depression would have averted 6,420 [1,475 – 11,366] person-years of extreme disability (i.e. HG_{G2}). In turn, the total avoidable disability, given the maximum HG, could reach 99,158 [89,661 – 108,837] person-years of extreme disability (i.e. $HG'_{G1} + HG'_{G2}$), and from that, 49,440 [42,529 – 56,370] are from people who are in treatment (i.e. HG'_{G2}; datum not shown in the table). Finally, 111,687 [95,775 – 128,599] are the person-years of extreme disability attributable to depression in people receiving treatment (i.e. DA'_{G2}).

Moreover, in accordance with the method mentioned previously, the values showed in Tables 4.4 and 4.5 can also be used to estimate the size of the healthcare intervention as a relative measure of effect. The relative benefit is equivalent to dividing 6,420 by 111,687 person-years of extreme disability (i.e. HG_{G2}/DA'_{G2}), suggesting that 5.8% [1.4 – 10.4] of the attributable disability in people receiving treatment is being avoided. Then, using Eq.3.2 from Chapter Three for the attributional fraction in the exposed population (i.e. 1-1/RR), the relative benefit can be transformed into a relative risk (RR) = 1.06 [1.01 – 1.10]. Furthermore, instead of using HG in the numerator, we can use the maximum HG (i.e. HG'_{G2}/DA'_{G2} = 49,440/111,687), and calculate a new RR = 1.44 [1.36 – 1.54].

Both relative risks are providing information about the effectiveness of treatment in terms of avoided disability. The first RR corresponds to the effect estimated using the average HG in people receiving the treatment (i.e. $RR_{HG\text{-}average}$). By contrast, the second corresponds to the estimated effect, assuming the maximum HG (i.e., $RR_{HG\text{-}max}$). According to our assumptions and the conceptual framework of effective coverage, the gap between both RRs would be a consequence of a deficit in the quality of the healthcare treatment for depression.

The explicit interpretation of the $RR_{HG\text{-}average}$ is that people receiving treatment are avoiding 1.06 times more person-years of extreme disability than if they were not receiving the treatment. This RR can also be expressed as a protective factor ($1/RR_{HG\text{-}average}$). In this case, the people receiving the treatment obtain 0.95 [0.90 – 0.99] of the person-years of extreme disability that they would obtain without receiving treatment. Expressing the results as RR allows comparison with other reports from the literature, where the effectiveness of healthcare interventions is traditionally shown by a relative measure of effect.

In this chapter, we have reviewed a formal proposal to measure effective coverage and several other related concepts. I introduced two different approaches for estimating them, and we have observed the results after their application to real data. In the next section, I

will discuss the options and assumptions that come from implementing these calculations. Additionally, the strengths and limitations of the overall proposal will be discussed.

4.6. Discussion

So far, I have provided formal definitions and expressions for effective coverage and other related indicators, producing consistent results. Whether using the population-average or the individual-level approaches, and whether based on model 1 or model 2, the results are apparently similar.

The advantage of the individual-level approach is that it allows us to obtain parameters of interpretability at the population level (such as person-years of extreme disability), as well as at the individual level (such as disability scores in the range between 0 and 100).

On the other hand, the advantage of using models that include interactions terms is that it allows us to generate a 'pseudo-variability' for the parameter HG, which means we can calculate an indicator for quality and effective coverage according to the framework provided by Shengelia et al.

4.6.1. Options and assumptions for implementing the proposal

Several decisions for implementing the proposal deserve more consideration, especially as they might be modified in future updates. These decisions are summarised in Table 4.6.

Table 4.6. Options and assumptions adopted, along with alternatives for implementing the estimation of effective coverage and other related concepts

N	Options and assumptions adopted in the process of implementation	Other alternatives
1	Estimates are produced using a single linear regression model.	Using at least two linear regression models, one for estimating the effect of normative need on disability, and another for estimating the effect of utilisation.
2	A linear regression model was used.	According to the features of the dependent variable (i.e., health/disability/utilities/quality of life outcome), another type of model could be used, such as a beta-regression model or a gamma-regression model.

3	The disability attributable (DA) and the health gain (HG) were defined as attribute from individuals, as well as from populations.	The population-average approach is dependent on population DA and HG, while the individual-level approach assumes that DA and HG is an individual-level attribute, aggregable to a population level..
4	The maximum HG is obtained using the observed distribution of the HG calculated at the individual level.	A distribution of the HG could be fitted using central and variability parameters, and then choosing the maximum HG.
5	Simulations are based on the expected and predicted values of each component of the effective coverage formulation	Simulation could be produced on a fitted distribution of the random variables and their counterfactuals. This option would imply adjusting the equations Eq.4.1-Eq.4.10.
6	Overall, a Normal distribution of parameters was chosen for simulations.	Whether negative values of HG are considered inappropriate, other distribution can be implemented, especially by the population-average approach.
7	In the individual-level approach, negative values for HG are permitted.	Different constraints can be implemented in cases where negative values of HG are considered inappropriate: transforming negative values into zeros or shifting up the observed distribution to positive numbers.

The first element from Table 4.6, and probably the more problematic, is the overall approach selected to calculate the parameters required for estimating effective coverage. This approach was based on a single linear regression model, including a specific option for how best to incorporate the features of normative need and utilisation. Another alternative would be using two separate linear regression models: one for assessing the effect of normative need on disability, and another to evaluate the effect of utilisation. The latter should be applied on a subsample of the data.

However, using one single linear regression model has the advantage of preserving the statistical power, even when the utilisation of healthcare is included as a covariate. In addition, having only one expression that summarises all the assumptions between variables makes it easier to implement in terms of coding a syntaxis for statistical software.

Despite these advantages, using one linear regression model might make the epidemiological and causal reasoning about what variables should be included more difficult. The interpretation of the resulting regression coefficients could also be less intuitive.

However, in the procedure developed, the main output of the linear regression models is the predicted values of disability. Therefore, the convenience of using one or two linear regression models should be assessed regarding their efficiency for making predictions. A systematic assessment of both alternatives is still pending; however, no significant differences were found in a preliminary assessment.

Another aspect related to the regression models (point 2, Table 4.6) is the assumption made about residuals. Even though the dependent variables ranged between 0 and 100, a linear regression model that assumes normality in the residual distribution was chosen. This decision was based on the fact that the variable of disability presented a distribution close to normal (see chapter two), and all primary estimations (D_{G0}, D_{G1} and D_{G2}, and their counterfactuals) took values far from the extremes of the disability's distribution. The decision was also made because disability might be replaced by other health state measures that are not necessarily restricted to fixed values. For instance, estimations using health dis-utilities (i.e. 1- utilities) can adopt values between zero and infinity. An exploration into the convenience of other regression models, for example, beta or gamma regression models, is still pending.

A third relevant element is to define whether the disability attributable (DA) and the health gain (HG) can be considered attributes of the individuals and/or from populations (point 3, Table 4.6). Using the population-average approach, DA and HG are calculated through population parameters and simulated according to their uncertainties. By contrast, in the individual-level approach, DA and HG are calculated at the individual level, to calculate their expected values or the expected totals. Then, these parameters are simulated according to their uncertainty, which implies that the DA and HG can be interpreted both as a population and an individual attribute, depending on which approach we chose.

However, it is important to note that the relative benefit, r-EC, a-EC, quality, and effective coverage are always calculated using simulated values from population parameters. This marks a difference from Shengelia et al.'s framework. It is also important to observe that the relative benefit and quality could have been calculated at the individual level, but that would not be consistent with the expression presented previously (i.e. Eq.4.1 – Eq.4.10).

A fourth important element to discuss, specific to the individual-level approach, is that the maximum HG is obtained from the individual predicted values of HG, from which we choose a certain threshold (i.e. the 0.9 percentile). However, another option for obtaining the maximum HG is to fit HG distribution using its estimated mean and standard deviation, including the covariance between both parameters, and then identify the value corresponding to the selected threshold. The second alternative is a little more complex; however, it might be conceptually more appropriate. In addition, the maximum HG was assumed to be a fixed value in the current implementation, even though the standard error for a quantile can be calculated and used to generate uncertainty around this parameter. Including the maximum HG uncertainty and fitting a distribution for the HG should be explored in more detail in future research.

A fifth element (point 5, Table 4.6), related to the previous one, is that when the data is simulated, we simulate values for the parameter of the expected or predicted level of disability, which agrees with equations Eq.4.1 – Eq.4.10. However, another option that might be explored in the future consists of simulating a population distribution of the random variable and its counterfactuals. When the expected or predicted values are estimated, we can obtain not just the mean and its standard error, but also the standard deviation including its own standard error, and the covariance matrix between the estimated mean and standard deviation. This option implies modifying the equations that describe effective coverage and other related concepts, including integrals for the distributions. How this alternative would affect the final estimate must still be investigated.

Another aspect to discuss (point 6, Table 4.6) is the underlying assumptions in the distributions selected to simulate the parameters' values, especially for the HG. In the population-average approach, normal distribution for simulating HG values was assumed (according to Table 4.4, HG= 0.8 [-2.0 to 3.6]), which considers the plausibility of obtaining negative values. Negative HG means assuming that healthcare interventions could be harmful. As a consequence of this option, the uncertainty intervals for the relative benefits (e.g. 5.9% [-18.3 to 26.7]), r-EC, and a-EC also included negative values. However, if the plausibility of negative values for HGs were constrained, assuming a gamma distribution for example, the results would have been very similar to those already received, though without negative values in the uncertainty intervals. For instance, the relative benefit estimation would have been 6.0% [0.0 to 38.0]. These constraints can be performed using a gamma

distribution only if the mean of HG is positive; otherwise, another procedure should be defined.

A seventh aspect of the procedure that deserves special and careful consideration is whether or not to constrain certain values, such as negative values in the HG. It is generally accepted that healthcare services can produce harm and, in some cases, more harm than good [25, 26]. In addition, several theoretical frameworks consider the side or collateral effects of health service interventions as a matter of quality. Therefore, if the procedure developed here can capture the adverse outcomes from the healthcare intervention, it should be welcomed, and negative values of HG should not be constrained.

However, in the case of the population-average approach, the negative values of the uncertainty intervals do not represent negative outcomes at the individual level, but an overall negative expected value of HG for the treatment of depression. To accept that the average effect of treatment for depression carries more disability than less is problematic, and therefore constraining negative values might be considered appropriate.

On the other hand, through the individual-level approach based on Model 2, an observed HG distribution can be obtained using the values calculated for each person. Figure 4.4 depicts such an observed distribution, with the position of the percentile 0.90 used as a maximum HG. It is noticeable that a negative HG was calculated for a great part of the population with treatment. To better understand those people with negative HG, in Figure 4.5 the HG is presented according to different variables.

Figure 4.4. Distribution of the health gains (HG) calculated for people with a depressive episode during the previous 12 months who are also receiving treatment. Data from the Chilean National Health Survey 2009-2020 (n=354).

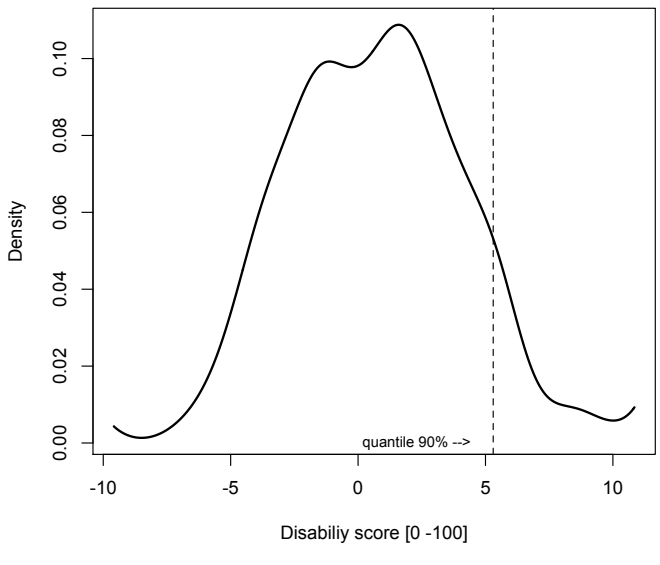
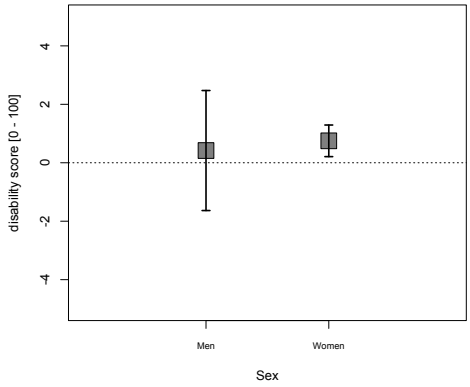
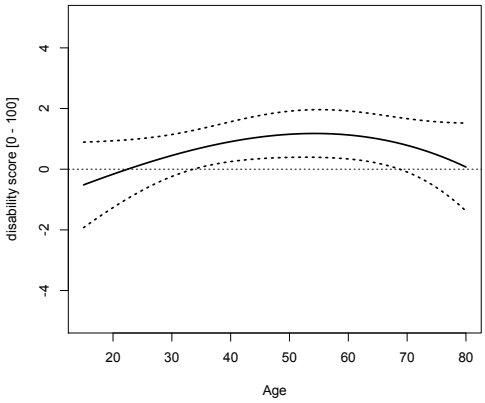


Figure 4.5. Health gains (HG) calculated for people with a depressive episode during the previous 12 months who are also receiving treatment, according to different variables. Data from the Chilean National Health Survey 2009-2020 (n=354).

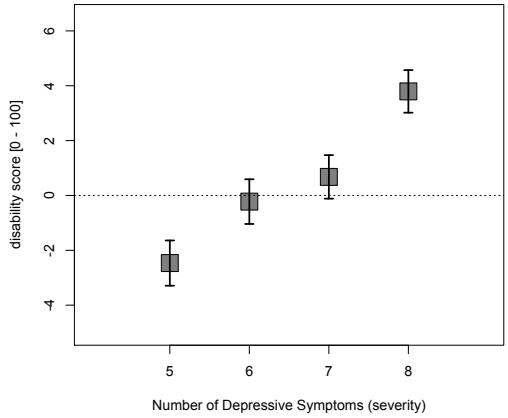
Panel A. HG by sex



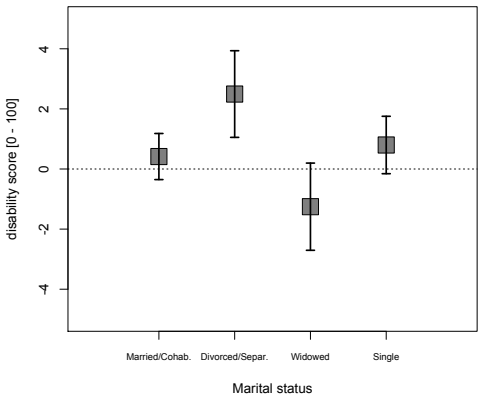
Panel B. HG by age



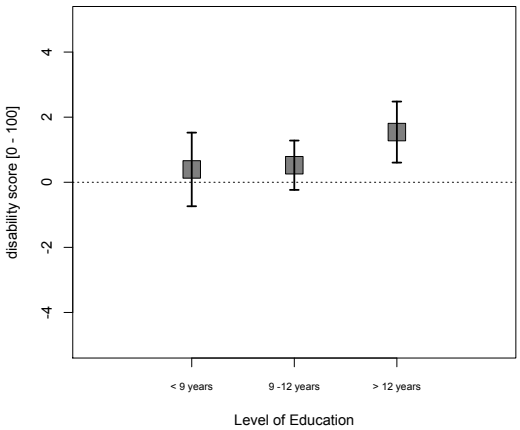
Panel C. HG by severity of depression*



Panel D. HG by marital status



Panel E. HG by level of education



Panel F. HG by quintile of income

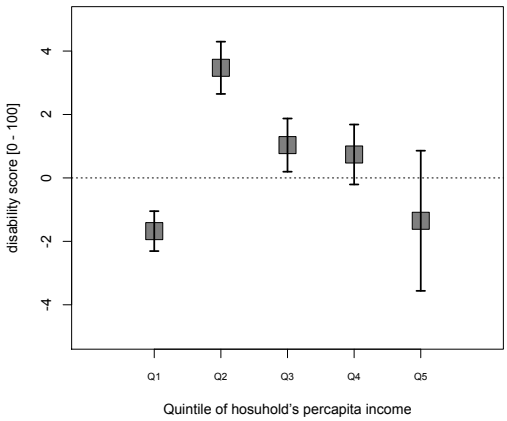
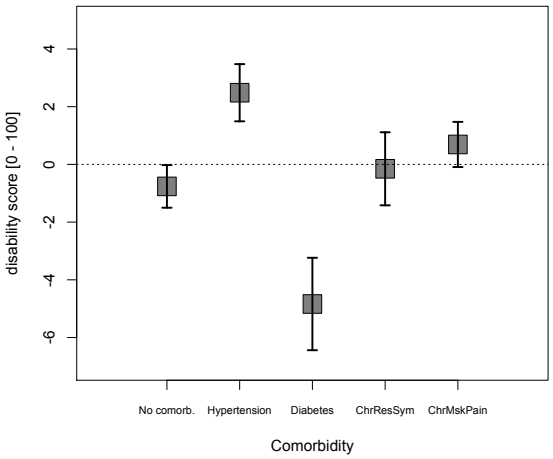


Figure 4.5. Health Gain (HG) calculated for people with a depressive episode during the previous 12 months who are also receiving treatment, according to different variables. Data from the Chilean National Health Survey 2009-2020 (n=354). (Continuing from previous page)

Panel G. HG by comorbidity



** People with fewer than five symptoms are not included among those with depression.*

No comob.: no comorbidity/ ChrResSym.: Chronic respiratory Symptoms/ ChrMskPain: Chronic musculoskeletal Pain

Interestingly, negative HG values are clearly observed in people with less severe depression, widowers, those from the poorest quintile of income, and people with diabetes. Is it reasonable to believe that people with these characteristics, on average, can obtain a negative benefit from the treatment of depression? For instance, people from the poorest quintile could be treated by low-quality providers, or maybe show poor adherence to the treatment. The evidence also shows a small or null effect of pharmacological treatment for mild cases of depression [27]. However, the possibility of an average harmful effect of depression treatment is still unlikely and remains a matter of discussion [28].

Excessive negative values of HG could indicate that the estimations from the linear regression model are likely biased, and residual confounding could remain after the adjustment by covariables. In this regard, it is important to remember that people's level of disability with the use of healthcare was higher than the observed disability in people without access to treatment for depression (see Table 4.2). Moreover, the variable of depression severity was introduced into the regression model to attempt to adjust the effect of utilisation by basal

conditions of disability. The topics of residual confusion and the problem of endogeneity will be further developed in chapter six.

One alternative for dealing with excessive negative HG values is to assume that all negative values result from a miss-specification of the model and imputing them as null HG (HG = 0). Another alternative is to assume that the bias is identical for all the individuals and, therefore, shifts up the whole observed distribution toward positive values. Table 4.7 presents the main results using the original approximation (equivalent to results shown in Table 4.5; an individual-level approach based on regression Model 2), and the results using both of the constraints already mentioned.

Table 4.7. Main results of effective coverage for depression, including different constraints on the health gains (HG) calculation, compared with the original results extracted from Table 4.5 (individual-level approach, model 2). Data from the Chilean National Health Survey 2009-2020 (n=4.447).

	Without constraining HG		Negative HGs are assumed equal to zero		Distribution of HG shifted up positive numbers	
	Individual-level approach		Individual-level approach		Individual-level approach	
	model 2		model 2		model 2	
	mean	**UI**	**mean**	**UI**	**mean**	**UI**
HG_{G2}	0.7	[-5.9 to 7.4]	1.8	[0.0 to 7.4]	10.3	[3.7 to 17]
$HG_{G2}max$	5.3	-	5.3	-	14.9	-
Relative benefit (%)	5.8	[1.4 - 10.4]	14.3	[10.5 - 18.5]	84.3	[68.7 - 102.3]
Quality (%)	13.1	[3.1 - 23.6]	32.3	[23.7 - 41.8]	68.9	[56.1 - 84.2]
Relative - effective coverage (%)	2.9	[0.7 - 5.2]	7.2	[5.4 - 9.1]	42.4	[35.3 - 50.2]
Absolute - effective coverage (%)	0.2	[0.0 - 0.3]	0.5	[0.4 - 0.6]	2.9	[2.5 - 3.3]
Effective coverage (%)	6.5	[1.6 - 11.6]	16.1	[12.0 - 20.4]	34.2	[28.6 - 40.2]
$HG'_{G1} + HG'_{G2}$	99159	[89661 – 108837]	99102	[89430 - 108885]	274592	[248391 - 300827]
HG_{G2}	6420	[1475 - 11365]	15878	[12093 - 19662]	93675	[80054 - 107296]
$RR_{HG\text{-}average}$	1.06	[1.01 - 1.10]	1.14	[1.10 - 1.19]	1.84	[1.69 - 2.02]
$RR_{HG\text{-}max}$	1.44	[1.36 - 1.54]	1.44	[1.36 - 1.54]	2.22	[2.00 - 2.50]

UI: uncertainty intervals

The first constraint produces more conservative results than the second. The maximum HG is kept unaltered (i.e. 5.3) but the HG increases (i.e. 1.8 disability units and 15,878 person-years of extreme disability), which produces higher relative benefit, quality, and effective coverage. The second alternative of constraint modifies the HG, as well as the maximum HG,

which modifies the maximum effectiveness of the healthcare intervention we can expect to achieve ($RR_{HG\text{-}max}$). The values of relative benefit and quality rise considerably.

It is important to note that the three effective coverage measures have uncertainty limits that do not overlap with each other. This means that, in practice, the impact of negative HG values is significant.

Apart from including constraints or trying new parameters for HG (see above), another option for leading with negative HG values is to change the overall method for estimating the potential benefit of the healthcare intervention in the context of data from cross-sectional studies. For instance, this could be done by using instrumental variables, propensity score matching, discontinuity regression models, or a more exhaustive strategy for prediction through machine-learning procedures. This subject will be developed in chapter six.

4.6.2. Limitations

The main limitation of the proposed procedure was mentioned in the previous paragraph and partially discussed in Chapter Three. The proposed procedure measures effective coverage by estimating the fraction of disability avoided by a healthcare intervention using cross-sectional data. The overall approach is grounded in linear regressions and is a natural extension of the developments presented in chapters Two (health-state valuations) and Three (burden of disability). Even though regression models are used for predictions, more than for causal analysis, the estimate of the benefit of the healthcare intervention is greatly dependent on the regression coefficient for utilisation.

The current level of disability can be correlated with the level of disability before starting the treatment. And, that initial level could have determined the likelihood of access to possible treatments, by different mechanisms to severity of symptoms or any other variable used to adjust the regression model. Therefore, the risk that residual confounding is biasing the outcomes cannot be completely discarded. Moreover, the presence of a large number of negative values of HG – in the individual-level approach – probably indicates this phenomenon.

Another limitation is the mismatch in the definitions between normative need and healthcare utilisation. The measurement of the normative need and the utilisation depends on how the survey items are formulated and how exhaustively the utilisation is explored. In this case, the 'ever treated' category is a broad measurement of use and is not restricted to the same temporal frame as the measurement of the normative need. In addition, the normative need is not measured with a level of specificity that allows different kinds of treatments to be investigated.

The correspondence between the measurement of the normative need and the utilisation is relevant. For instance, had the survey inquired about the use of a specific type of talking therapy, it would have also been necessary to identify those cases of depression with the specific normative need for that intervention.

By contrast, when the utilisation is broadly formulated, as in the case of the Ch-NHS 2009-2010, it assumes that the same level of benefit from the healthcare intervention can be expected in all cases. In the current analysis, the procedure can predict the HG according to different severity levels; however, the maximum HG is chosen by assuming that all individuals could reach that unique level of benefit. These elements can be addressed by using lower values of the threshold for selecting the maximum HG. However, the threshold remains essentially arbitrary.

Four sensitivity scenarios using different definitions of normative need and utilisation were also presented. Complete results are given in Tables S4.2-6 from the supplementary material. However, to discuss the implication of varying the normative need and utilisation definitions, only those results for effective coverage are shown in Table 4.8, including the constraint of negative values in the HG.

According to the first sensitivity scenario, to coincide with the temporal frameworks for the normative need (i.e. 12 months) and the treatment items, the cases of depression were restricted only to those examples where the first episode occurred during the last two years. Under that scenario, the prevalence of the normative need diminished dramatically to only 3.6%, but the utilisation increased to 74.6%. Despite this, the effective coverage remains relatively constant, and the uncertainty intervals overlap with the basal scenario results.

When I change the definition of utilization to 'medication during the last two weeks' (i.e. the second sensitivity scenario), the utilisation falls to 19.5%, and the effective coverage assumes negative values. This can be explained because several cases that received treatment but are not using medication during the last two weeks were included among those without healthcare utilisation. Using constraints to negative HG values, negative values for effective coverage are also avoided.

Table 4.8. Effective coverage for depression according to different definitions of normative need, utilisation and constraints on the health gains (HG) calculation. Individual-level approach, model 2. Data from the Chilean National Health Survey 2009-2020 (n=4.447).

	Basal scenario*		**Senario 2**		**Scenario 3**		**Scenario 4**	
NN defintion	All cases		First episode during last 2 years		All cases		Severe cases	
Ut definition	Ever treated		Ever treated		Medication in last 2 weeks		Medication in last 2 weeks	
NN (prevalence) (%)	18.1	[16.1 - 20.1]	3.6	[2.6 - 4.5]	18.1	[16.1 - 20.1]	4.0	[2.9 - 5.1]
Ut (coverage) (%)	49.6	[43.6 - 55.6]	74.6	[63.2 - 86.1]	19.5	[14.8 - 24.2]	22.1	[10.8 - 33.3]
Effective Coverage (%) (without constrains in HG)	6.5	[1.6 to 11.6]	4.3	[-10.7 to 19.4]	-9.8	[-15.5 to -4.3]	-3.1	[-8.7 to 2.5]
Effective Coverage (%) (Negative HG assumed zero)	16.1	[12.0 to 20.4]	23.0	[11.6 to 36.1]	4.2	[1.6 to 6.9]	3.9	[1.0 to 7.1]

*NN: normative need/ Ut: utilization/ *Values from the basal scenario are the same showed in Table 4.6.*

In the last sensitivity scenario, the analysis restricted the sample only to those with more severe symptoms, where it is assumed that medication is always required. The normative need diminishes, but the utilisation remains relatively constant. In this new scenario, the effective coverage remains at low levels.

In summary, the results seem to vary depending on how the normative need and the utilisation are defined, and how both parameters relate to each other.

Despite these limitations, it is important to notice that the main results are not far from other estimates. Data from the World Mental Health Survey, using a large sample from 21 countries, estimated that the fraction of people with a 12-month major depressive disorder who received a 'minimally adequate treatment' – defined as ≥ 1 month of medication plus ≥ 4 visits to any type of medical doctor, or ≥ 8 visits with any professional – was only 16.5%, with a raw coverage of 40.1% [22, 29]. In the present analysis, we obtained a raw coverage

of 49.6% and effective coverage (using health gains) equivalent to 6.5%, which rises to 16.1% when we constrain negative values of HG.

Another piece of evidence about the plausibility of the results comes from a recent network meta-analysis that pooled information from more than one thousand participants from randomised control trials attempting to estimate the effectiveness of pharmacological interventions for depression [30]. For example, when compared with a placebo, fluoxetine showed a pooled OR equal to 1.52 [1.40 – 1.66]. Most of the studies in the meta-analysis measured the response to fluoxetine using scales of symptoms, and efficacy was defined by the reduction of ≥50% in the standardised score. On the other hand, the results from this chapter support an RR_{HGmax} equal to 1.44 [1.36 – 1.54], using disability for the outcome. Though the two measures are not directly comparable, the scale of the results did not seem very different.

Finally, another limitation inherent to surveys is that diseases with low frequency or short duration usually cannot be investigated properly. This restricts the scope of healthcare interventions that can be covered by the developed procedure to measure effective coverage. However, this makes surveys especially suitable for chronic non-communicable diseases. It is worth remembering that, according to the systematic review from chapter one, only 26.4% of the research on effective coverage was conducted in non-communicable diseases and mental health disorders.

4.6.3. Methodological strengths of the proposal

The procedure presented in this chapter has several advantages compared with current literature that tries to measure effective coverage.

First, the proposed procedure is consistent with the framework used by Shengelia et al. for effective coverage. It includes the component of health gain, which is a major difference from the existing literature on the topic. The absence of adequate health gain and quality parameters is the main shortcoming of current measurements of effective coverage [31].

Moreover, the procedure is theoretically underpinned in the subjects of health valuation, attributable fractions, and burden of disease – all of them with strong roots in the field of epidemiology and health metrics.

Second, the procedure expands the repertoire of indicators used to assess the impact and performance of healthcare interventions. Under this procedure, not only is the effective coverage calculated, but also the relative benefit, r-EC, a-EC, $RR_{averageHG}$, RR_{maxHG}, and the quality of the healthcare intervention. The utility of each of these indicators will be explored in the next chapter, when the procedure is applied to compare the effective coverage between diseases and countries.

Moreover, the results of the procedure are expressed through different metric units, such as those from the individual level (e.g. scores of disability) or those from the population level (e.g. person-years of extreme disability).

Third, the procedure can be implemented on different outcomes used to measure the heath gain, either according to the availability within surveys or according to the preferred conceptual framework. Throughout the last three chapters, I have used the construct of disability, but the procedure can be adapted for other measures of health, health utilities, or quality of life. Health utilities are especially interesting when resource allocation and decision-making are relevant. On the other hand, when the outcome is anchored in reference values – such as full health and death – the results might also be combined with mortality outcomes.

Fourth, all the estimates are based on a single source of information: a national health survey from which the normative need, utilisation, and health gain are all extracted. This ensures internal consistency between parameters, as well as external consistency concerning the target population.

Fifth, the methodology is based on regression models, a procedure familiar to most public health scientists. Regression models also offer a flexible framework to discuss and test what variables (and assumptions) should be included in the analysis. Moreover, since the steps taken to produce the estimates are clearly defined, it is fairly simple to implement through a function for statistical software. Having a function for statistical software, such as the one presented in the supplementary material (see Main functions for R S.4.1), allows us to

explore the effect of adding new confounders or interaction terms. What's more, hypothetical scenarios can be created using pseudo-data on covariables, which is useful to assess their impact on the current level of effective coverage. The graphics presented in Figure 4.4. are only an example of the automatic outputs that the function can offer for measuring effective coverage.

When it comes to making estimates, this procedure offers further advantages. For example, the outputs can easily be combined with other health services measures, such as the budgets addressed to healthcare interventions, thresholds of cost-effectiveness, or the current levels of the burden of disease. A discussion of these measures will be given in chapter six.

4.7. Conclusion

In this chapter, I have reviewed the rationality and formal expressions for effective coverage and a set of related indicators; I discussed how to implement these expressions practically; I tested the procedure using real data; and I discussed the results according to different approaches. Special emphasis was given to the transparency of all the assumptions that underpin the proposal, and alternatives for future implementation were provided. Finally, the strengths and limitations of the procedure were presented.

In summary, I have created a proposal to estimate effective coverage – including the health gain attributable to healthcare interventions – in the context of cross-sectional data. Although this proposal seems attractive due to the diversity of indicators it can generate, it is not free from limitations. Moreover, its performance, reliability, and utility still need to be explored in more depth, applying it to different contexts (countries) and healthcare interventions using a different source of information.

References of chapter four

1. Shengelia B, Tandon A, B. Adams O, Murray CJL: **Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy**. *Social Science & Medicine* 2005, **61**:97-109.
2. World Health Organization: **Tracking universal health coverage: first global monitoring report**. In.; 2015.
3. World Health Organization and the International Bank for Reconstruction and Development / The World Bank: **Tracking universal health coverage: 2017 global monitoring report**. Switzerland.; 2017.
4. Ng M, Fullman N, Dieleman JL, Flaxman AD, Murray CJL, Lim SS: **Effective Coverage: A Metric for Monitoring Universal Health Coverage**. *PLOS Medicine* 2015, **11**(9).
5. Richard C, Roy S, Caroline G: **Jonathan Bradshaw on Social Policy. Selecting Writings 1972 - 2011**. York, United Kingdom: University of York; 2013.
6. KJ Rothman, S G, TL L: **Modern epidemiology**. Philadelphia: Lippincott Williams & Wilkins; 2008.
7. Ouellet RP, Apostolides AY, Entwisle G, Hebel JR: **Estimated community impact of hypertension control in a high risk population**. *Am J Epidemiol* 1979, **109**(5):531-538.
8. Lalonde M: **A New Perspective on the Health of Canadians, a working document**. In. Ottawa: Governemment of Canada; 1974.
9. Malhi GS, Mann JJ: **Depression**. *The Lancet* 2018, **392**(10161):2299-2312.
10. Angst J, Gamma A, Rössler W, Ajdacic V, Klein DN: **Long-term depression versus episodic major depression: Results from the prospective Zurich study of a community sample**. *Journal of Affective Disorders* 2009, **115**(1):112-121.
11. Markkula N, Zitko P, Pena S, Margozzini P, Retamal CP: **Prevalence, trends, correlates and treatment of depression in Chile in 2003 to 2010**. *Soc Psychiatry Psychiatr Epidemiol* 2017, **52**(4):399-409.
12. C. KR, Gavin A, Daniel M, Bedirhan U, Hans-Ulrich W: **The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF)**. *International Journal of Methods in Psychiatric Research* 1998, **7**(4):171-185.
13. Association AP: **Diagnostic and Statistical Manual of Mental Disorders**. Washington, DC: American Psychiatric Association; 1994.
14. Health NCCfM: **Depression. The NICE Guideline on the Treatment and Management of Depression in Adults. Updated Edition 2020**. United Kingdom: The British Psychological Society & The Royal College of Psychiatrists; 2010.
15. Rai D, Zitko P, Jones K, Lynch J, Araya R: **Country- and individual-level socioeconomic determinants of depression: multilevel cross-national comparison**. *Br J Psychiatry* 2013, **202**(3):195-203.
16. Andrews G, Henderson S, Hall W: **Prevalence, comorbidity, disability and service utilisation. Overview of the Australian National Mental Health Survey**. *Br J Psychiatry* 2001, **178**:145-153.

17. 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**. *The Lancet* 2007, **370**(9590):851-858.
18. Cornelius LR, van der Klink JJ, Groothoff JW, Brouwer S: **Prognostic factors of long term disability due to mental disorders: a systematic review**. *J Occup Rehabil* 2011, **21**(2):259-274.
19. Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, Adeyi O, Barker P, Daelmans B, Doubova SV *et al*: **High-quality health systems in the Sustainable Development Goals era: time for a revolution**. *The Lancet Global health* 2018, **6**(11):e1196-e1252.
20. Ayuso-Mateos JL, Nuevo R, Verdes E, Naidoo N, Chatterji S: **From depressive symptoms to depressive disorders: the relevance of thresholds**. *British Journal of Psychiatry* 2010, **196**(5):365-371.
21. Mitchell AJ, Malone D, Doebbeling CC: **Quality of medical care for people with and without comorbid mental illness and substance misuse: systematic review of comparative studies**. *Br J Psychiatry* 2009, **194**(6):491-499.
22. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, Bruffaerts R, de Girolamo G, de Graaf R, Gureje O *et al*: **Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys**. *Lancet* 2007, **370**(9590):841-850.
23. Greenland S, Pearce N: **Statistical foundations for model-based adjustments**. *Annu Rev Public Health* 2015, **36**:89-108.
24. Palazzo C, Yokota RTC, Ferguson J, Tafforeau J, Ravaud JF, Van Oyen H, Nusselder WJ: **Methods to assess the contribution of diseases to disability using cross-sectional studies: comparison of different versions of the attributable fraction and the attribution method**. *Int J Epidemiol* 2018.
25. Peer RF, Shabir N: **Iatrogenesis: A review on nature, extent, and distribution of healthcare hazards**. *J Family Med Prim Care* 2018, **7**(2):309-314.
26. Lunevicius R, Haagsma JA: **Incidence and mortality from adverse effects of medical treatment in the UK, 1990–2013: levels, trends, patterns and comparisons**. *International Journal for Quality in Health Care* 2018, **30**(7):558-564.
27. Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, Fawcett J: **Antidepressant drug effects and depression severity: a patient-level meta-analysis**. *Jama* 2010, **303**(1):47-53.
28. McGauran N, Wieseler B, Kreis J, Schüler Y-B, Kölsch H, Kaiser T: **Reporting bias in medical research - a narrative review**. *Trials* 2010, **11**(1):37.
29. Thornicroft G, Chatterji S, Evans-Lacko S, Gruber M, Sampson N, Aguilar-Gaxiola S, Al-Hamzawi A, Alonso J, Andrade L, Borges G *et al*: **Undertreatment of people with major depressive disorder in 21 countries**. *Br J Psychiatry* 2017, **210**(2):119-124.
30. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JPT *et al*: **Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis**. *The Lancet* 2018, **391**(10128):1357-1366.

31. Jannati A, Sadeghi V, Imani A, Saadati M: **Effective coverage as a new approach to health system performance assessment: a scoping review**. *BMC health services research* 2018, **18**(1):886.
32. Lim SS, Collaborators aG: **Measuring the health-related Sustainable Development Goals in 188 countries: a baseline analysis from the Global Burden of Disease Study 2015**. *The Lancet* 2016, **8**(10053):1813-1850.

Chapter 5: Estimating effective coverage for several health conditions across different countries

5.1. Introduction

Effective coverage has been proposed as an indicator for tracking progress towards the goal of Universal Health Coverage. Several institutions have supported the metric proposed by Shengelia et al. [1] as the most convenient and appropriate [2, 3] to assess effective coverage. However, this approach is flawed because there is no measure to calculate the health gains associated with healthcare interventions in each context where effective coverage is measured [4, 5].

In the first chapter of this thesis, through a scoping review, I demonstrated the lack of studies in the scientific literature that addressed this problem when measuring effective coverage.

In the second chapter, I introduced the concepts of health and health state valuation in order to achieve a measure that incorporated health gains associated with healthcare interventions. Additionally, I argued that through a latent variable calculated using a health state description questionnaire, it is possible to estimate the disability and health weights attributable to health conditions.

In chapter three, using this approach and applying concepts related to attributable fractions, I presented a proposal to calculate the burden of disability attributable to a health condition using cross-sectional data.

Based on this approach, in chapter four, I presented a proposal to calculate effective coverage, including the health gains of a given healthcare intervention. Essentially, the proposal consisted of estimating the fraction of the avoidable disability attributable to a disease avoided by a healthcare intervention. In that chapter, the concept of effective

coverage was also opened to several other indicators that describe different healthcare provision aspects, such as the relative benefit of the intervention, the quality of the intervention, and the absolute and relative effective coverages (a-EC and r-EC, respectively).

However, all these developments have been tested using a single database from a national health survey carried out in a single country. What's more, the performance of the proposed effective coverage estimate was evaluated on a single healthcare intervention.

Chapter five aims to assess the results of applying this proposal to several countries and different healthcare interventions.

In the next section of this chapter, I will introduce a new database, the selected health conditions and healthcare interventions, and some specific methodological considerations. Intermediate results, such as the psychometric properties of the used disability questionnaire, will also be provided in that section.

In sections three to six, I will present the effective coverage results for each specific healthcare intervention, explaining and discussing the main findings of the proposal's performance, including limitations (i.e. negative health gains estimates).

The seventh section of this chapter will contain a short discussion of the proposal and an overall view of using different effective coverage indicators to assess various aspects of healthcare performance.

5.2. Data, variables, definitions, and statistical methods

5.2.1. Data and sampling

To evaluate the proposal's performance in measuring effective coverage, I used data from the World Health Organisation (WHO) study on global ageing and adult health (SAGE). This cohort study was carried out in six lower- and upper-middle-income countries using representative samples of the general population. The countries included were: China, Ghana, India, Mexico, the Russian Federation, and South Africa. The overall aim of the study was to provide comparable information on health and wellbeing, as well as to track the impact of health interventions and policies within and across countries [6].

Specifically, I used the data from Wave 1, undertaken between 2007 and 2010. The sample included people older than 17 years, with an overrepresentation of those aged 50 years or more. Except for Mexico, all samples were obtained using multistage clustering strategies. In Mexico, the sample included supplementary and replacement samples for losses in the follow-up. Sample weights were calculated considering the sampling strategy and a post-stratification factor according to the age and sex distribution within each country.

The interviews were conducted face to face and included a lengthy questionnaire that asked about health state, disability, socioeconomic factors, risk factors, the presence of diseases, and use of healthcare services, among other topics. Additionally, the interviews considered anthropometric measures such as blood pressure and spirometry. More information is available elsewhere [6] (see: https://www.who.int/healthinfo/sage/en/; consulted in January 2021). The response rate varied between 93% in China to 53% in Mexico.

This study was chosen because it gathers together several elements that make it convenient for this thesis. Firstly, it provides information from several countries, diseases and healthcare interventions. Secondly, the study also includes a variety of questionnaires for measuring health state and related concepts, as well as confounding variables for the relationships between disability-diseases and disability-utilisation of healthcare interventions – all requirements for measuring effective coverage. Thirdly, because this study was led by an internationally reputable organisation (WHO), it guarantees a certain quality in the surveys process and the comparability of procedures across countries. Additionally, being part of the WHO, the data is freely accessible. Fourthly, sample sizes were relatively large, with Wave 1

consisting of over 2,500 individuals in each country. Finally, the participating countries belong to different world regions, making the comparison of results more generalisable.

5.2.2. Variables and definitions

According to the proposal presented in chapter four, in order to measure effective coverage, four kinds of information are required: a normative need; a healthcare intervention addressing that need; an outcome on a continuous scale such as health status, disability, health state utilities, or health-related quality of life; and potential confounders that could be used to adjust the models.

The questionnaire of the SAGE study, Wave 1, includes items for identifying several health conditions. However, a relatively small number had specific items exploring coverage by healthcare interventions. Throughout the questionnaire, four normative needs were identified with their respective items about the utilisation of healthcare: depressive disorder, hypertension, osteoarthritis, and chronic obstructive pulmonary disease (COPD). From these four, COPD was dismissed because its definition was based on spirometric parameters, and in some countries its measurement accounted for numerous missing and implausible values.

5.2.2.1. Depression

For depressive disorder, the SAGE study included a modified version of the Composite International Diagnostic Interview used in the World Mental Health Survey [7]. This version, consisting of eighteen items, asks about the presence of sadness, emptiness, depression, loss of interest, or decreased energy during the previous twelve months, lasting for most of the day over two weeks or more, along with several other symptoms. I applied the DSM-IV criteria for identifying cases with a depressive disorder [8]. The utilisation of the healthcare intervention addressing depressive disorders was defined according to the question: 'Have you been taking any medications or other treatment for it during the last 12 months?'. In this way, it is assumed that all cases of depression during the last 12 months need 'medication or other treatment'. Notice that the normative need and the utilisation are measured during the same time interval (i.e. 12 months).

5.2.2.2. Hypertension

Hypertension was defined as either systolic blood pressure higher than 140 mmHg, or diastolic blood pressure higher than 90 mmHg, or both, using the lowest value from three readings. The blood pressure measurements were taken with the participant seated. Participants with normal blood pressure, but who declared that they were taking 'medications or other treatment' for this condition during the last two weeks, were also assumed to be suffering from hypertension [9, 10]. Consequently, the healthcare intervention was defined as 'medications or other treatment' for 'hypertension or high blood pressure', and the utilisation was considered during the last two weeks.

5.2.2.3. Osteoarthritis

Osteoarthritis was defined as the presence, during the last 12 months, of 'pain, aching, stiffness or swelling in or around the joints (like arms, hands, legs or feet) which was not related to an injury and lasted for more than a month', as well as 'stiffness in the joint in the morning after getting up from bed, or after a long rest of the joint without movement'. This stiffness lasted for 30 minutes or less, and went away 'after exercise or movement in the joint'. Moreover, the symptoms should have been present during the last two weeks. On the other hand, healthcare utilisation was again defined as receiving 'medications or other treatment' for this condition in the previous two weeks. It is assumed that all cases with osteoarthritis need healthcare intervention.

5.2.2.4. Health–disability outcome

To quantify the potential health gain associated with a healthcare intervention, I used the Health State Description questionnaire. This consisted of twenty items exploring eight domains of functioning (mobility, self-care, pain and discomfort, cognition, interpersonal activities, sleep and energy, affect, and vision), plus two items about overall health. Each answer has five Likert alternatives, ranging from no problem to extreme problems in activity or participation (see chapter two for full definitions). This instrument is very similar to the one used in chapters two, three and four of this thesis, and belongs to the same family of questionnaires used by the WHO in other studies [11] (see chapter two). Also similar to

chapter two, the questionnaire's psychometric properties were assessed for each country separately.

Briefly, a description of each item in the questionnaire was carried out (see supplementary tables S5.1-6), and a polychoric correlation matrix was explored in half of the SAGE sample (see supplementary figures S5.1-6). The number of factors was calculated through a parallel analysis [12], applied to the polychoric correlation matrix using half of the sample selected randomly in each country (see results in supplementary figures S5.7-12). Then, an exploratory factor analysis (EFA; weighted least square method, oblimin rotation) was performed in half of the SAGE sample, exploring multidimensionality (see tables S5.7-12 in supplementary material). Thirdly, a confirmatory factor analysis (CFA) for categorical data was carried out on the other half of the sample, exploring: a) a unidimensional solution, b) a second-order model according to the EFA, c) a second-order model according to the theoretical framework of the questionnaire, consisting of nine first-order latent variables, one for each domain of disability, including the overall health domain, and a second-order latent variable [13]. The solutions were compared through a Satorra et al. test [14]. Path diagrams of the best factorial solution for each country are presented in the supplementary material (see Figures S5.12-18).

In summary, the questionnaire proved adequate to its theoretical structure in Ghana and the Russian Federation. However, the EFA's solution showed better goodness of fit in China, India, Mexico and South Africa. In China, India and Mexico, items for mobility were gathered in a single domain jointly with the items of general health; while in South Africa, the items for mobility, general health and self-care were accounted for by a unique first-order latent variable. Different parameters for the goodness of fit, by country, applied to the entire sample are shown in Table 5.1. A detailed description of these procedures is available in chapter two.

Table 5.1. Model fit and indices of the confirmatory factor analysis of the selected model for the Health State Description questionnaire- The SAGE Study, Wave 1, analysis by country.

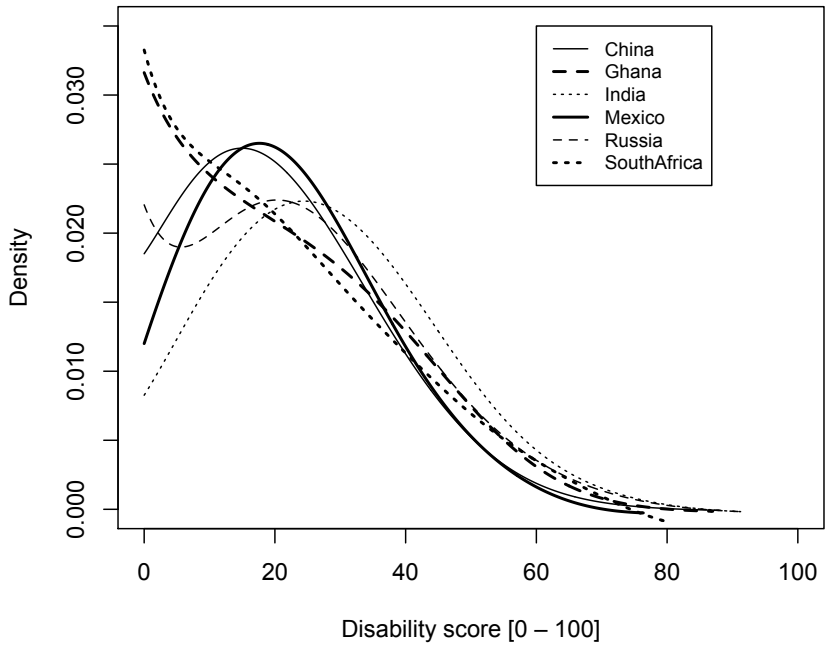
	China	Ghana	India	Mexico	Federation Russia	South Africa
n	14231	5015	11229	2636	4339	4146
χ^2	6932.2	7536.0	10019.8	1699.8	4060.8	4761.9
Df	201	200	201	201	200	202
p. value	0.000	0.000	0.000	0.000	0.000	0.000
TLI	1.00	0.99	0.99	0.99	1.00	1.00
CFI	1.00	0.99	1.00	0.99	1.00	1.00
GFI	1.00	0.99	1.00	0.99	1.00	1.00
RMSEA	0.05	0.09	0.07	0.05	0.07	0.08
UCI90%	0.05	0.08	0.07	0.05	0.07	0.07
LCI90%	0.05	0.09	0.07	0.06	0.07	0.08
SRMR	0.05	0.07	0.05	0.05	0.06	0.06

χ^2: chi-square / Df: degree of freedom / TLI: Tucker-Lewis index / CFI: comparative fir index / GFI: goodness of fit index / RMSEA: root mean square error of approximation / SRMR: standardised root mean square residual / LCI: lower 90% confidence interval / UCI: upper 90% confidence interval.
Note: Cut-off criterion for an adequate goodness of fit are: TLI, CFI and GFI ≥ 0.95; RMSE ≤ 0.06; and SRMR ≤ 0.08 [15]

The reliability of the questionnaire was evaluated using Cronbach's alpha [16], showing a good level of internal consistency, between 0.92 [0.91 – 0.92] in Mexico and 0.95 [0.94 – 0.95] in South Africa.

The second-order latent variable, corresponding to the construct of disability, was calculated and standardised on a scale between 0 and 100, where '0' means the absence of disability and 100 means extreme disability in all the questionnaire items. Figure 5.1. presents the distribution of this variable for each country.

Figure 5.1. Distribution of the latent variable of disability, from the Health State questionnaire, by country. SAGE study, Wave 1.



From this, we can see that the distributions are overlapped, especially in values higher than 30 units of disability.

5.2.2.5. Confounding factors

Finally, other variables included in the analysis of effective coverage that might be potential confounders were included, such as sex, age, zone (two categories: rural or urban), marital status (four categories: currently married, never married, separated/divorced, and widowed), education (four categories: less than primary, primary completed, secondary completed, and college/university completed), quintile of income within each country, and occupation status (three categories: working, never worked/homemaker, and not working). The diseases described above were also assumed to be potential confounding factors.

5.2.3. Statistical methods

A description of the sample according to the variables included in the study was performed for each country. The relationship between disability and each variable was explored by a multivariable linear regression model separately in each country.

The prevalence and the coverage of healthcare interventions for depressive disorder, hypertension and osteoarthritis were reported. The burden of disability attributable to each health condition was calculated through the individual-level approach described in chapter three. The procedure was applied using the same multivariable linear regression models described in the previous paragraph. The analysis of the burden of disability was implemented to confirm consistency with the effective coverage results. The option of implementing a multilevel approach (e.g. a random-effect model), is discussed in the last section of this chapter.

Similar to the burden of disability, the effective coverage for each healthcare intervention was explored using a multivariable linear regression model separately for each country. The regression models included variables identifying people with the normative need but without utilisation (G1), and people with the normative need and with utilisation (G2).

The initial regression model included interaction terms between G1, G2 and all other covariables. However, the regression models used for the effective coverage calculation included only those parameters chosen after a backward selection process (see chapter four for details) [17]. In the case of depressive disorder, a variable of severity was added to the model, built using the number of depressive symptoms (i.e. between 0 and 9). Throughout the questionnaire, it was not possible to find markers of severity for hypertension and osteoarthritis. The effective coverage was implemented using the individual-level approach described in chapter four. In that chapter, it was argued that results from the individual-level and the population-average approaches were equivalent. The 90th quantile was used as a threshold to choose a maximum health gain for all normative needs in all the countries.

In a sensitivity analysis, the results were presented, including restriction to negative values for the parameter of health gain. A second sensitivity analysis was also performed, assuming effective coverage as an individual attribute (see explanation in Section 5.4).

All procedures were carried out considering the sample weights and were implemented using the statistical software R 3.5.0. A confidence interval of 95%, or uncertainty intervals corresponding to quantiles 2.5% and 97.5%, are reported according to the procedure of estimation.

5.3. General results

From the 42,489 people interviewed, 38,086 (89.6%) had complete data in all the variables included in the analysis. The highest percentage of missing data occurred in South Africa (27.7%), while the smallest occurred in India (3.4%). No data imputation was carried out because the main aim of this study was to evaluate the performance of a procedure rather than generate estimates. In addition, the overall proportion of missing data was relatively low – less than 20%. A more detailed analysis of missing data is available in the supplementary Table S5.13.

The description of the sample is shown in Table 5.2. The mean age was similar across countries, while the mean disability score was higher in India (27.5 [27.1 – 28.0]) and smaller in South Africa (21.8 [18.6 - 24.9]). The proportion of women was about 50%. Rurality varied considerably between countries, being higher in India (74.7% [73.2 - 76.3]) and lower in Mexico (18.1% [14.1 - 22.0]). The stratum of currently married accounted for the most significant proportion in all countries. An education level less than primary was especially elevated in Ghana and India, while the Russian Federation showed the highest proportion of people with complete secondary or further education. Regarding occupation status, the most-answered category was ‘working’, especially in Ghana (81.2% [78.9 - 83.4]).

Table 5.2. Description of the sample by country. SAGE study, Wave 1.

	China (n=13,311)			Ghana (n=4,819)			India (n=10,846)		
		mean	CI		mean	CI		mean	CI
Age	-	45.6	[45.1 - 46]	-	44.3	[43.7 - 45]	-	41.0	[40,6 – 41.5]
Disability	-	21.1	[20.6 - 21.6]	-	22.2	[21.3 - 23.1]	-	27.5	[27.1 – 28.0]
	N	**%**	**CI**	**N**	**%**	**CI**	**N**	**%**	**CI**
Sex (women)	7143	48.6	[46.4 - 50.9]	2289	49.8	[46.8 - 52.9]	6663	49.2	[47.6 - 50.9]
Zone (rural)	6846	51.9	[49.7 - 54.2]	2843	53.6	[50.5 - 56.7]	8109	74.7	[73.2 – 76.3]
Marital status (currently married)	11143	89.3	[88.1 - 90.6]	2881	72.5	[69.9 - 75.1]	8429	82.1	[80.8 - 83.3]
Marital status (never married)	233	5.4	[4.3 - 6.5]	134	8.4	[6.5 - 10.3]	600	9.3	[8.2 - 10.4]
Marital status (separated/divorced)	240	1.6	[1.1 - 2]	638	9.0	[7.3 - 10.6]	74	0.5	[0.3 - 0.7]
Marital status (widowed)	1695	3.7	[3.3 - 4.2]	1166	10.1	[8.7 - 11.6]	1743	8.2	[7.5 - 8.9]
Education (less than primary)	5289	19.4	[18 - 20.9]	2949	43.5	[40.5 - 46.5]	6034	44.7	[43.1 - 46.3]
Education (completed primary)	2625	18.8	[17 - 20.6]	603	19.6	[17 - 22.1]	1659	16.6	[15.3 - 17.8]
Education (completed secondary)	4753	53.3	[51 - 55.5]	1093	32.4	[29.4 - 35.3]	2517	30.4	[28.7 - 32.1]
Education (completed college/university)	644	8.5	[7.2 - 9.8]	174	4.6	[3.2 - 6]	636	8.3	[7.2 - 9.4]
Income quintile 1 (poorer)	2588	10.1	[9.1 - 11.1]	938	15.4	[13.2 - 17.5]	1935	20.7	[19.3 - 22.1]
Income quintile 2	2691	16.4	[14.9 - 17.9]	922	16.6	[14.4 - 18.8]	2080	21.2	[19.8 - 22.6]
Income quintile 3	2686	19.1	[17.4 - 20.9]	967	19.3	[16.9 - 21.7]	2080	19.9	[18.5 - 21.2]
Income quintile 4	2747	23.8	[21.8 - 25.7]	1003	22.7	[20.1 - 25.3]	2281	18.1	[16.8 - 19.3]
Income quintile 5 (richer)	2599	30.6	[28.3 - 32.9]	989	26.1	[23.2 - 28.9]	2470	20.2	[18.8 - 21.5]
Occupation (working)	5768	67.7	[65.7 - 69.7]	3467	81.2	[78.9 - 83.4]	4613	56.3	[54.7 - 57.9]
Occupation (never working/homemaker)	710	4.6	[3.7 - 5.5]	58	1.6	[0.7 - 2.5]	2477	19.0	[17.9 - 20.1]
Occupation (not working)	6833	27.7	[25.8 - 29.5]	1294	17.2	[15.2 - 19.3]	3756	24.7	[23.4 - 26]

CI Confidence Interval 95%

Table 5.2. Description of the sample by country. SAGE study, Wave 1. (Continuation from the previous page)

	Mexico (n=2,451)			Russian Federation (n=3,603)			South Africa (n=3,056)		
		mean	CI		mean	CI		mean	CI
Age	-	42.7	[40.8 - 44.5]	-	47.0	[44.7 - 49.4]	-	41.9	[40.1 – 43.7]
Disability	-	21.8	[20.3 - 23.4]	-	24.3	[21.7 - 26.9]	-	21.8	[18.6 – 24.9]
	N	**%**	**CI**	**N**	**%**	**CI**	**N**	**%**	**CI**
Sex (women)	1523	52.4	[44.6 - 60.2]	2323	55.2	[47.4 - 62.9]	1847	55.0	[46.5 - 63.5]
Zone (rural)	665	23.3	[17.6 - 29.1]	887	18.1	[14.1 – 22.0]	1015	26.6	[20.5 – 32.7]
Marital status (currently married)	1560	70.6	[63.4 - 77.9]	2039	60.2	[53 - 67.4]	1550	50.0	[41.5 - 58.5]
Marital status (never married)	234	20.3	[13 - 27.6]	157	13.3	[7.7 - 18.8]	482	30.9	[23.2 - 38.6]
Marital status (separated/divorced)	148	4.3	[2.4 - 6.3]	315	13.7	[8.2 - 19.1]	208	9.8	[4.1 - 15.4]
Marital status (widowed)	509	4.7	[3.4 - 6.1]	1092	12.9	[10.4 - 15.3]	816	9.3	[6.8 - 11.8]
Education (less than primary)	1329	26.5	[20.8 - 32.1]	105	0.9	[0.6 - 1.2]	1486	21.9	[16.5 - 27.4]
Education (completed primary)	544	25.5	[17.9 - 33.2]	260	2.2	[1.6 - 2.8]	702	13.9	[10.2 - 17.7]
Education (completed secondary)	358	35.3	[27.7 - 43]	2532	76.1	[70.1 - 82.2]	700	56.2	[48.3 - 64.1]
Education (completed college/university)	220	12.7	[7.8 - 17.5]	706	20.8	[14.8 - 26.8]	168	8.0	[4.0 – 12.0]
Income quintile 1 (poorer)	508	15.3	[10.8 - 19.8]	661	12.5	[8.3 - 16.6]	585	18.9	[13.8 - 24.1]
Income quintile 2	506	24.3	[15.4 - 33.2]	701	12.6	[9.7 - 15.5]	615	19.0	[12.0 – 26.0]
Income quintile 3	454	20.9	[15 - 26.8]	716	16.2	[10.5 - 21.9]	619	23.5	[15 - 32.1]
Income quintile 4	502	14.2	[10.6 - 17.9]	738	24.2	[17.7 - 30.7]	630	19.8	[13.3 - 26.4]
Income quintile 5 (richer)	481	25.3	[19.3 - 31.2]	787	34.5	[26.5 - 42.5]	607	18.7	[12.7 - 24.7]
Occupation (working)	740	52.3	[44.6 - 59.9]	1351	65.2	[58.7 - 71.8]	852	43.0	[34.1 - 51.9]
Occupation (never working/homemaker)	951	27.4	[21.9 - 32.9]	62	1.9	[0.7 - 3.1]	318	6.4	[4.2 - 8.6]
Occupation (not working)	760	20.3	[13.9 - 26.7]	2190	32.8	[26.4 - 39.2]	1886	50.6	[42.1 - 59.2]

CI Confidence Interval 95%

The relationship between disability and all the other variables under study was investigated using multivariable regression models performed separately in each country, presented in Table 5.3. Age was associated with disability. Women showed higher levels of disability than men in all countries, except in the Russian Federation and South Africa, where the difference was not significant. South Africa was the only country where living in a rural area was associated with less disability. There was no association between marital status and disability. In general, in China, India, Mexico and the Russian Federation, those with higher education levels showed less disability. In Ghana and South Africa, education was not associated with disability score. Regarding the quintile of income, all countries showed less disability in higher-income quintiles. Overall, the category of not working, compared with the working category, was associated with higher disability.

Depressive disorder was the health condition with the greatest level of disability, but with wide differences between countries. For instance, people with depression in China showed 20.3 [15.3 – 25.2] more additional units of disability than people without depression, while in Ghana this magnitude reached only 5.7 [2.4 – 8.9] units. Osteoarthritis, in all the countries, was the health condition with the second-highest associated disability. People with hypertension showed a small increase in disability compared to people without hypertension. The regression coefficient from hypertension showed no statistical significance in Mexico, the Russian Federation and South Africa.

Table 5.3 Regression coefficients (mean differences) for disability (multivariate analysis), according to different variables, by country. SAGE study, Wave 1.

	n= 13,311		n= 4,819		n= 10,846		n= 2,451		n= 3603		n= 3,056	
	China		Ghana		india		Mexico		Russian Federation		South Africa	
	Coef	**CI**	**Coef**	**CI**	**Coef**	**CI**	**Coef**	**CI**	**Coef**	**CI**	**Coef**	**CI**
(Intercept)	6.7	[3.7 - 9.7]	0.0	[-3.8 - 3.8]	8.4	[6.3 - 10.5]	15.7	[9.3 - 22]	9.5	[1.4 - 17.5]	0.6	[-9.4 - 10.7]
Age (by 10 years)	3.6	[3.2 - 4.0]	3.9	[3.4 - 4.4]	4.0	[3.7 - 4.3]	2.2	[1.4 - 3.1]	4.2	[3.3 - 5.1]	4.2	[2.9 - 5.4]
Sex (women)	1.5	[0.6 - 2.4]	1.9	[0.2 - 3.6]	3.7	[2.8 - 4.7]	4.2	[1.8 - 6.7]	-0.4	[-3.7 - 2.9]	-0.3	[-6.2 - 5.6]
Zone (rural)	2.2	[1.1 - 3.3]	3.2	[1.5 - 4.8]	1.8	[0.8 - 2.8]	1.7	[-1.3 - 4.6]	4.4	[1.6 - 7.1]	-5.3	[-10.4 - -0.3]
Marital status (currently married)	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-
Marital status (never married)	2.0	[-0.8 - 4.8]	4.1	[0.8 - 7.4]	-0.6	[-2.1 - 0.9]	-0.2	[-3.0 - 2.6]	0.4	[-4.3 - 5.1]	3.3	[-3.5 - 10.2]
Marital status (separated/divorced)	-0.6	[-3.7 - 2.4]	1.5	[-1.3 - 4.4]	0.6	[-3.7 - 4.9]	3.7	[-0.6 - 8.0]	2.3	[-1.4 - 6.0]	1.5	[-5.0 - 8.1]
Marital status (widowed)	0.3	[-0.9 - 1.5]	3.8	[1.5 - 6.0]	1.3	[0.1 - 2.5]	-0.9	[-4.7 - 2.9]	0.9	[-1.7 - 3.5]	0.0	[-4.5 - 4.6]
Education (less than primary)	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-
Education (completed primary)	-1.2	[-2.6 - 0.2]	-0.8	[-3.0 - 1.5]	-0.4	[-1.5 - 0.8]	-2.3	[-5.6 – 1.0]	-6.7	[-11.7 - -1.7]	-1.6	[-6.4 - 3.1]
Education (completed secondary)	-2.7	[-4.0 - -1.4]	-0.7	[-2.7 - 1.2]	-2.3	[-3.3 - -1.2]	-4.6	[-7.8 - -1.5]	-9.0	[-13.8 - -4.3]	-0.8	[-6.3 - 4.6]
Education (completed college/university)	-3.2	[-5.6 - -0.9]	0.4	[-3.4 - 4.1]	-4.7	[-6.4 - -3.0]	-3.3	[-7.4 - 0.8]	-8.4	[-14.1 - -2.7]	1.1	[-5.2 - 7.3]
Income quintile 1 (poorer)	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-
Income quintile 2	-2.9	[-4.5 - -1.3]	1.0	[-1.5 - 3.5]	-0.6	[-1.9 - 0.6]	-3.6	[-8 - 0.7]	1.2	[-1.9 - 4.4]	0.0	[-5.9 - 6.0]
Income quintile 3	-3.8	[-5.4 - -2.2]	1.1	[-1.5 - 3.7]	-1.8	[-3.0 - -0.6]	-3.8	[-7.7 - 0.2]	0.3	[-2.6 - 3.3]	-2.1	[-10.3 - 6.2]
Income quintile 4	-5.1	[-6.6 - -3.6]	-2.7	[-5.3 - -0.1]	-2.8	[-4.1 - -1.5]	-6.5	[-10.2 - -2.9]	-2.6	[-6.5 - 1.3]	-3.9	[-8.7 - 0.9]
Income quintile 5 (richer)	-6.5	[-8.0 - -4.9]	-1.8	[-4.5 - 0.8]	-4.8	[-6.2 - -3.5]	-7.1	[-11 - -3.1]	-1.9	[-5.7 - 2.0]	-6.1	[-12 - -0.1]
Occupation (working)	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-
Occupation (never working/homemaker)	2.1	[0.1 - 4.0]	-3.4	[-7.4 - 0.6]	0.9	[-0.1 - 1.9]	-0.8	[-3.6 - 2.1]	4.1	[-3.5 - 11.8]	6.6	[0.5 - 12.8]
Occupation (not working)	4.6	[3.3 - 5.8]	6.8	[4.7 - 8.8]	2.1	[1.1 - 3.1]	2.0	[-0.7 - 4.7]	6.3	[4.1 - 8.4]	6.3	[1.4 - 11.2]
Depression	20.3	[15.3 - 25.2]	5.7	[2.4 - 8.9]	12.4	[10.9 - 14]	10.1	[6.2 - 13.9]	12.5	[9.2 - 15.8]	16.2	[8.9 - 23.4]
Hypertension	0.6	[-0.4 - 1.5]	0.1	[-1.5 - 1.7]	1.9	[0.9 - 3.0]	-1.0	[-3.4 - 1.3]	1.6	[-2.2 - 5.5]	3.6	[-1.1 - 8.3]
Osteoarthritis	6.8	[5.2 - 8.5]	11.8	[9.5 - 14.2]	6.6	[5.3 - 7.8]	6.6	[2.3 - 11.0]	5.6	[3.4 - 7.9]	9.5	[4.9 - 14.1]

Coef: regression coefficients/ CI: Confidence Interval 95%

5.4. Effective coverage of the healthcare addressing depressive disorders

5.4.1. Description of the main results

The prevalence of depressive disorder observed in the six countries under study ranged between 0.9% in China and 7.8% in India (see Table 5.4) consistent with data from other studies [7, 18]. The burden of disability attributable to depression, which integrates prevalence and disability in a single metric, varied between 0.8% (China) and 3.6% (South Africa). These values are similar to other estimations based on years lived with disability [IHME: https://vizhub.healthdata.org/gbd-compare/ consulted in February 2021 for the year 2007], except for India and South Africa, where the values presented in Table 5.4 are higher.

The coverage of 'medication or other treatments' varied significantly across different countries. India showed the smallest coverage, as only 5.5% of people with depression reported treatment during the last 12 months. On the contrary, in South Africa, the coverage reached up to 75.9%. The observed values are similar to those in previous reports [19].

The coefficients from regression models used to calculate the effective coverage in each country are presented in Table S5.14 of the supplementary material, while Tables 5.4 and 5.5 show the main results from the analysis.

To help interpret the results of effective coverage, which were widely discussed in chapter four, I will briefly describe the results for China and then discuss the procedure used to estimate effective coverage. In China, from a sample representative of 270 million people (see Table 5.5), depressive disorders would have been responsible for 472,699 person-years of extreme disability in the survey year. By contrast, treatment coverage would have avoided 13,204 extra person-years of extreme disability (see Table 5.5). The relative benefit was estimated equal to 13.2% (≈ 13,204/69,406; see Tables 5.4 and 5.5), which means that nearly a sixth of the disability attributable to depression is avoided by the healthcare intervention among those who are receiving treatment. This relative benefit is equivalent to a relative risk of 1.16 (i.e. 1/(1-relative benefit)).

On the other hand, assuming that all people receiving treatment achieve at least the maximum health gain, the total avoided disability would have been 14,690 person-years of extreme disability. Since the avoided disability in this counterfactual scenario and the actual

scenario are similar (i.e. 14,690 is close to 13,204), the estimated quality of the healthcare is considered relatively high (66.8%), and the maximum RR is close to the RR from the relative benefit: 1.17.

Because the relative benefit is low (only 13.2%), and the raw coverage is also small (i.e. 14.3%), the relative effect coverage (r-EC) is just 2.7%, which means that only 2.7% of the disability attributable to depression in China is being avoided. When this benefit is displaced to the total population disability, the absolute effective coverage (a-EC) reaches 0.02%. Finally, the effective coverage, which is conceptually equivalent to the product of the coverage and the quality (see chapter four), is 9.7%.

Table 5.4. Burden of disability, prevalence, coverage of the treatment and results from the effective coverage analysis of the healthcare for depressive disorder, by country. SAGE study, Wave 1.

	n=13,311		n=4,819		n=10,846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Burden (%)	0.8	[0.5 - 1.2]	1.1	[0.8 - 1.4]	3.5	[3.2 - 3.8]
Prevalence (%)	0.9	[0.4 - 1.3]	4.3	[3.1 - 5.5]	7.8	[7.0 - 8.6]
Coverage (%)	14.3	[-7.8 - 36.5]	7.5	[-1.1 - 16.1]	5.5	[3.0- 8.0]
Relative benefit (%)	13.2	[-188.9 - 218.8]	-49.0	[-985.1 - 829.6]	-20.9	[-65.7 - 6.7]
Quality (%)	66.8	[-799.8 - 984.8]	-101.6	[-1043.9 - 452.7]	-47.1	[-124.3 - 15.4]
Relative - effective coverage (%)	2.7	[-2.3 - 9.4]	-10.3	[-38.2 - 13.3]	-1.4	[-3.3 - 0.4]
Absolute - effective coverage (%)	0.02	[0 - 0.1]	-0.11	[-0.3 - 0.1]	-0.05	[-0.1 - 0.0]
Effective coverage (%)	9.7	[-8.1 - 33.4]	-8.8	[-29 - 10.7]	-2.4	[-5.6 - 0.8]
RR HG-average	1.16	[0.42 - Inf]	0.69	[0.1 - Inf]	0.84	[0.62 - 1.1]
RR HG-maximum	1.17	[0.34 - Inf]	1.85	[0.13 - Inf]	1.79	[1.33 - 11.12]
	n=2,451		n=3,603		n=3,056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Burden (%)	2.8	[1.6 - 4.0]	1.3	[0.9 - 1.28]	3.6	[0.2 - 7.3]
Prevalence (%)	6.0	[3.4 - 8.6]	2.6	[1.6 - 3.6]	4.9	[0.2 - 9.6]
Coverage (%)	24.5	[8.8 - 40.2]	24.6	[7.9 - 41.3]	75.9	[50.0 - 101.7]
Relative benefit (%)	10.3	[-51.6 - 80.4]	29.7	[-1.1 - 140.5]	30.3	[-314.4 - 423.7]
Quality (%)	11.4	[-56.7 - 87.7]	64.5	[1.6 - 250.3]	47.4	[-329.1 - 498.4]
Relative - effective coverage (%)	2.2	[-10.3 - 16]	10.4	[0.6 - 23.6]	27.7	[-184.5 - 278.7]
Absolute - effective coverage (%)	0.07	[-0.3 - 0.4]	0.16	[0.0 - 0.3]	1.77	[-0.3 - 3.9]
Effective coverage (%)	2.6	[-11.9 - 17.9]	15.3	[0.9 - 32.8]	38.5	[-19.9 - 263.7]
RR HG-average	1.08	[0.66 - 5.32]	1.42	[0.99 - Inf]	1.45	[0.29 - Inf]
RR HG-maximum	8.98	[1.43 - Inf]	1.84	[1.09 - Inf]	2.31	[0.16 - Inf]

CI: confidence or creditable intervals/ HG: health gain/ RR: relative risk/ HG: health gain

Table 5.5. Disability attributable to depressive disorder according to the analysis of burden and results from the effective coverage analysis of healthcare for depressive disorder expressed in person-years. SAGE study, Wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
n weighted	270,697,618	-	11,052,687	-	279,003,445	-
Disability Attributable (burden analysis)	472,699	[250595 - 693240]	27,029	[20076 - 34143]	2,701,820	[2458311 - 2949450]
DA_{G1}	422,061	[197252 - 646870]	22,640	[13081 - 32198]	2,516,213	[2222654 - 2809772]
DA'_{G2}	69,406	[-49613 - 188426]	3,631	[-3119 - 10382]	178,459	[81338 - 275581]
HG_{G2}	13,204	[-10370 - 36778]	-2,664	[-8484 - 3157]	-37,678	[-87797 - 12441]
HG'_{G1}	120,881	[54809 - 186952]	28,309	[20862 - 35757]	1,472,232	[1294413 - 1650051]
HG'_{G2}	14,690	[-8456 - 37837]	2,461	[-207 - 5129]	79,696	[49964 - 109429]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
n weighted	53,571	-	83,442,881	-	21,452,752	-
Disability Attributable (burden analysis)	323	[192 - 456]	272,390	[185383 - 357792]	170,480	[2164 - 334394]
DA_{G1}	290	[115 - 466]	205,240	[126280 - 284200]	49,816	[12790 - 86843]
DA'_{G2}	79	[29 - 130]	107,901	[20363 - 195440]	218,432	[-109309 - 546172]
HG_{G2}	8	[-33 - 48]	32,428	[2334 - 62522]	84,442	[-14318 - 183201]
HG'_{G1}	242	[106 - 378]	162,359	[101697 - 223021]	55,532	[23087 - 87977]
HG'_{G2}	71	[29 - 114]	50,289	[12776 - 87803]	156,573	[-41536 - 354683]

DA_{G1}: disability attributable in people with depression without treatment (G1)/ DA'_{G2}: disability attributable in people with depression and treatment (G2) assuming they are not receiving the treatment/ HG_{G2}: health gain from people in G2/ HG'_{G1}: maximum health gain from G1/ HG'_{G2}: maximum health gain from G2.

5.4.2. Central estimations

Regarding the general performance of the procedure to calculate effective coverage, it seems to have worked relatively well using data from China, Mexico, the Russian Federation, and South Africa, at least in terms of the central estimations and the plausibility of results.

However, in the case of Ghana and India, we can observe in Table 5.5 that the central estimate of the health gain was negative, which scarcely seems possible. As a consequence of negative health gains, all the estimates that used this measure in any part of their calculation also produced negative results (see Table 5.4). Figure 5.2 shows the density curves for each country's health gain (see supplementary Figure S5.19 for separate

histograms). In that figure, it is observed that India and Ghana are the countries with more significant displacement of their density of health gain towards negative values.

Figure 5.2. Observed distribution of the health gain attributable to healthcare interventions for depressive disorder, by country. SAGE study, Wave 1.

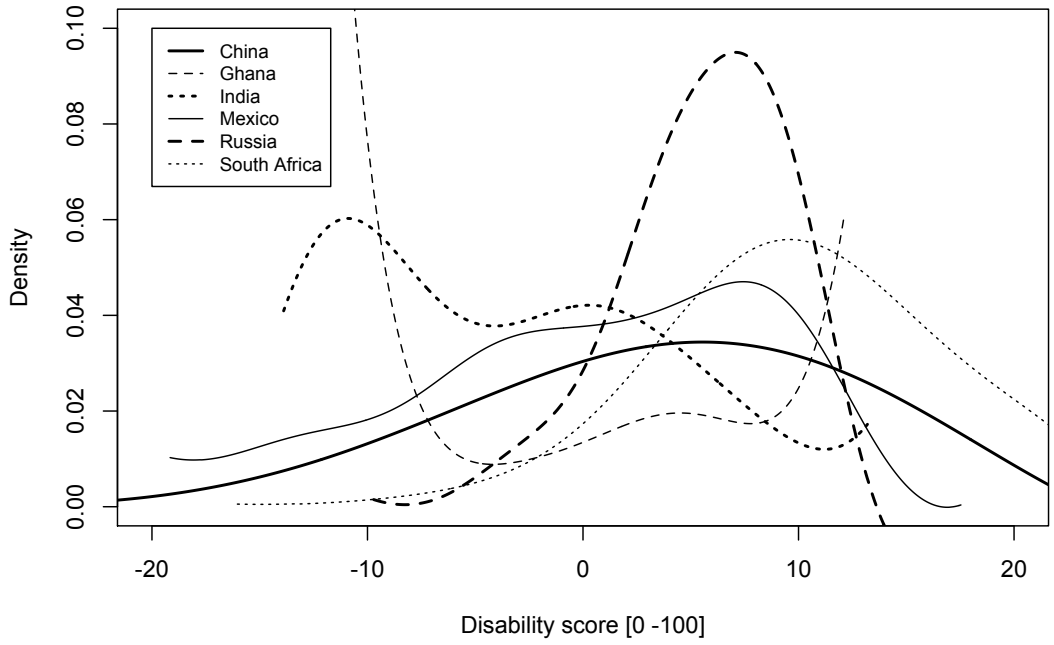


Table 5.6. presents the results of effective coverage, assuming negative health gain valued at zero (sensitivity analysis). We can observe that India and Ghana's results seem more plausible, while estimates for China, the Russian Federation and South Africa are almost unaltered. Mexico, however, a country with a relatively high proportion of negative health gain estimations, shows improvement in the indicators. The quality of healthcare, for example, rises in Mexico from 11.4% up to 43.9%, and the effective coverage from 2.6% up to 9.9%.

Table 5.6. Results of effective coverage of healthcare for depressive disorder, restricting negative values of health gain (sensitivity analysis). SAGE study, Wave 1.

	n=13,311		n=4,819		n=10,846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	14.5	[-178 - 221.9]	12.1	[-149.6 - 185]	11.3	[4.7 - 27.7]
Quality (%)	74.9	[-863.1 - 1091.8]	25.1	[-75.8 - 223]	25.4	[10.9 - 47.6]
Relative - effective coverage (%)	2.9	[-2.2 - 9.9]	2.5	[-0.6 - 6.8]	0.7	[0.3 - 1.2]
Absolute - effective coverage (%)	0.02	[0.0 - 0.1]	0.03	[0.0 - 0.1]	0.03	[0.0 – 0.0]
Effective coverage (%)	10.3	[-7.7 - 36.1]	2.2	[-0.5 - 5.1]	1.3	[0.6 – 2.0]
RR HG-average	1.16	[0.29 - Inf]	1.15	[0.32 - Inf]	1.13	[1.05 - 1.4]
RR HG-max	1.20	[0.35 - Inf]	1.89	[0.14 - Inf]	1.82	[1.31 - Inf]

	n=2,451		n=3,603		n=3,056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	39.5	[-3.4 - 136.6]	30.6	[0.8 - 140.8]	31.5	[-329.6 - 394.3]
Quality (%)	43.9	[-3.6 - 140.5]	66.3	[6.1 - 274.6]	47.8	[-327.9 - 473.2]
Relative - effective coverage (%)	8.5	[-0.6 - 23.5]	10.8	[1.2 - 23.9]	28.6	[-170.9 - 260.6]
Absolute - effective coverage (%)	0.27	[0.0 - 0.6]	0.16	[0.0 - 0.3]	1.82	[-0.3 - 4.0]
Effective coverage (%)	9.9	[-0.7 - 26]	15.7	[1.7 - 33.6]	38.9	[-23.0 - 256.9]
RR HG-average	1.64	[0.97 - Inf]	1.47	[1.0 - Inf]	1.49	[0.32 - Inf]
RR HG-max	7.17	[1.41 - Inf]	1.80	[1.07 - Inf]	2.41	[0.21 - Inf]

5.4.3. Uncertainty

The procedure tends to generate large uncertainty intervals, which is more evident in South Africa. This may be a consequence of the sample size, the variability of parameters, or the procedures. Regarding the sample size, all estimates from people with treatment are performed using small sample sizes. For example, in China, from a sample of 13,311 people, only 130 had a depressive disorder, and nine received treatment. This explains why the uncertainty intervals are larger in estimates from people with treatment (G2) than from people without treatment (G1), observed in Table 5.5. Small samples of people with treatment imply considerable uncertainty on the health gains attributable to the healthcare intervention, and since all the indicators of effective coverage include health gains in some part of their equations, these indicators will be affected by such uncertainty. This is especially relevant to relative benefit and quality indicators, since they are estimators exclusively based on people with treatment. On the contrary, smaller uncertainty intervals are observed in the indicator of a-EC, which combines information from the whole sample (i.e. people with depression (with or without treatment) and people without depression). The country with

the largest sample size from people with treatment was Mexico, but this still meant only 58 individuals.

From the perspective of the parameters' variability, we must remember that estimates in Table 5.5 come from individual predictions, which were later added, according to the individual-level approach described in chapter four. Thus, the variance of the effective coverage parameters depends on the regression models that generated such predictions, including the effect of the collinearity between variables. However, as mentioned in chapter four, the complexity of models was not an important source of variability.

A third cause of large uncertainty intervals can be attributed to how effective coverage indicators are conceived. All these indicators carry the division between health gains and other parameters. Therefore, when either the numerator or denominator is close to zero, extreme results are easily generated during the simulation process. This may be avoided if we estimate effective coverage for individuals, and then average the estimates to obtain a population estimator, as originally suggested by Shengelia et al. [1]. Indeed, the outputs generated by the procedures proposed in this thesis allow an easy calculation of the relative benefit, the quality, and the effective coverage for each individual on the database (second sensitivity analysis: see supplementary Table S5.15). The estimates, following this approach, are similar to those presented in Table 5.4, though in the case of relative benefit and quality, the uncertainty intervals are narrower. The width of uncertainty intervals of effective coverage remains almost unaltered.

However, the approach from Shengelia is not entirely coherent with the reasoning used to justify the approaches presented in chapter four. Further discussion about the differences between Shengelia's proposal and the proposal suggested in this thesis are discussed in chapter six, including the advantages and disadvantages of each one.

5.4.4. Coherence with the metric of burden of disability attributable

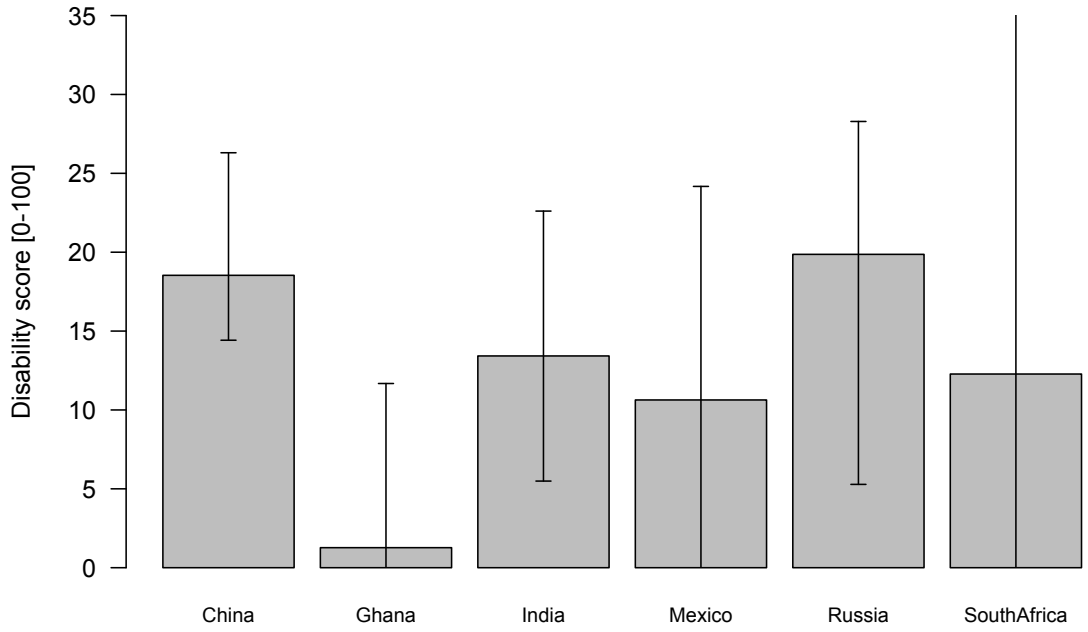
Another important finding of this thesis's proposed procedure is that the burden of disability attributable to depression, calculated using the approach described in chapter three, seems consistent with the estimates of disability attributable from the procedure used to calculate effective coverage. Table 5.5 shows that the disability attributable to depression calculated

by the burden analysis (the approach used in chapter three) approximately coincides with the sum of DA_{G1} and $DA'_{G2} - HG_{G2}$ (parameters from effective coverage analysis). This means the disability attributable to depression from people without treatment (DA_{G1}) plus the disability attributable from people with treatment (DA'_{G2}), discounting the effect of the treatment (HG_{G2}). The agreement confirms the metric coherence between burden and effective coverage.

5.4.5. Interpretability according to burden of disability attributable

Another aspect to bear in mind when assessing the procedure's performance is the potential to reduce disability that any healthcare intervention might have. For example, the disability attributable to depressive disorder in each country demonstrates this point (see Figure 5.3).

Figure. 5.3. Expected disability attributable to depressive disorder at the individual level, according to the procedure used to calculate effective coverage by country. SAGE study, Wave 1



Ghana, for example, showed the lowest disability attributable to depression, while the Russian Federation and China showed the largest, though with significant overlapping of

uncertainty intervals. This is relevant because, at a population level, in cases such as Ghana it seems that healthcare interventions have little to contribute in terms of reducing disability, or little to offer in terms of possible health gains, because cases with depression are only associated with a small amount of disability.

The relevance of this issue can be further explained. According to Table 5.6, the central estimation of the relative benefit for the treatment for depression is similar between Ghana (12.1%), China (14.5%) and India (13.2%). However, this relative benefit (i.e. the fraction of the attributable disability that is avoided) was calculated over very different levels of attributable disability, which represent different levels of absolute benefits. The absolute benefit can be calculated by dividing the number of person-years avoided from the healthcare intervention (see HG_{G2}, Table 5.5) by the number of people with depressive disorder under treatment.

5.4.6. Quality and effectiveness

Continuing with the comparison between countries, according to Table 5.6, China and the Russian Federation show a relatively high level and similar quality of treatment: 74.9% and 66.6%, respectively. However, the Russian Federation has a higher relative benefit (the intervention avoids a higher fraction of the attributable disability) than China: 30.6% versus 14.5%, respectively. This means that, despite similar quality, on average, the Russian Federation's intervention shows higher effectiveness. This can be noticed in the RR-average, which is another way of expressing the relative benefit, where the Russian Federation has a value of 1.47 and China 1.16. This inconsistency between quality and effectiveness is expected, since quality and effectiveness (either relative benefit or RR-average) are different concepts that provide complementary information about healthcare performance.

Moreover, quality is calculated using standards of comparison taken from within each country. In other words, in a country where everyone obtains the same health gain, even if it is small, the quality will be 100%, since everyone is achieving the maximum benefit observed in that population.

Finally, we observe that the RR-maximum – i.e. a relative expression of the fraction of the maximum health gain over the attributable disability – differed notably across countries. In

China, it was only 1.20, while in Mexico it reached 7.17, and, in the four other countries, it was about 2.0 (see Table 5.6). This indicator shows the maximum effectiveness achievable by the healthcare intervention in each country, while differences may be related to the health technology available in each place.

However, the RR-maximum must be considered with caution because, as mentioned above, it is calculated over the attributable disability. For example, Ghana and the Russian Federation had an RR-maximum equal to 1.89 and 1.80, respectively, representing around 45% of the attributable disability when we assume no healthcare intervention. However, the expected disability attributable to depressive disorder at the individual level was less than five units of disability in Ghana, but about 20 in the Russian Federation. Therefore, the maximum health gain is around 2.25 units of disability in Ghana, and 9.0 in the Russian Federation, which shows that, in absolute terms, the Russian Federation is offering a significantly higher maximum benefit (or maximum effectiveness, results do not show). This limitation, common in relative indicators, can be overcome by calculating a rate of effectiveness between the number of person-years avoided under the maximum benefit assumption (HG'G2, from Table 5.5) and the number of people with depressive disorder under treatment. Importantly, when the constraint of negative values in health gain is implemented (sensitivity analysis), the RR-maximum does not change largely (see Tables 5.6 and 5.4).

5.5. Effective coverage of the healthcare addressing hypertension

5.5.1. Main results: negative values in HG

The prevalence of hypertension by country is described in Table 5.7. India showed the smallest prevalence (15.5%) and South Africa the highest (46.8%). This range of values is coincident with other studies [20].

Regarding results for the burden of disability, the attributable disabilities calculated for hypertension were relatively small and close to zero – except for South Africa (see Table 5.3). However, given the large numbers of people with hypertension, the burden shows relatively high values. This is especially true for South Africa, which obtained the largest disability attributable and the highest prevalence, reaching a burden equal to 7.8% of the total disability in that country.

Mexico showed a negative burden. This result is counterintuitive, and it is a consequence of the disability attributable obtained for that country in the regression model used by the procedure (see Table 5.3). In other words, this means that people from that country with hypertension, on average, show a marginally lower level of disability (or higher health) than people without hypertension, regardless of all confounders included in the regression model.

The obtained results are difficult to compare, since the Institute of Health Metric and Evaluation (IHME) does not directly investigate the burden of hypertension. IHME only reports the burden of hypertensive heart disease, which is a type of heart failure due to hypertension, and whose burden – measured as years lived with disability – ranges between 0.045% and 0.32% in the countries under study. These values are smaller than most of the values of the burden of disability reported in Table 5.7 [IHME: https://vizhub.healthdata.org/gbd-compare/ consulted in February 2021 for the year 2007]. Additionally, the IHME reports the disability weight given for a general uncomplicated disease that requires daily medication as 0.049 (on a scale between 0 and 1) [21]. However, this value is higher than the disability attributable calculated through the regression models presented in Table 5.3. In chapter two, I argued that disability weights are conceptually comparable to the disability attributable.

Table 5.7. Burden of disability, prevalence, coverage of the treatment and results of effective coverage of healthcare for hypertension, by country. SAGE study, Wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Burden (%)	0.9	[0.8 - 1.0]	0.2	[0.1 - 0.2]	1.1	[1.0 - 1.2]
Prevalence (%)	33.7	[31.6 - 35.8]	34.9	[32 - 37.8]	15.5	[14.4 - 16.6]
Coverage (%)	24.4	[21.6 - 27.2]	12.7	[9.6 - 15.8]	23.5	[20.5 - 26.5]
Relative benefit (%)	-354.5	[-631.8 - -226]	-286.4	[-5084 - 4051.6]	-329.5	[-629.4 - -212.8]
Quality (%)	-92.2	[-117.1 - -69.9]	-37.6	[-74.1 - -6.5]	-148.8	[-185 - -117.7]
Relative - effective coverage (%)	-179.1	[-438.6 - -106]	85.5	[-1825 - 1781.2]	-47.5	[-62.9 - -36.7]
Absolute - effective coverage (%)	-0.9	[-1.1 - -0.7]	-0.5	[-0.9 - -0.1]	-0.3	[-0.4 - -0.3]
Effective coverage (%)	-23.2	[-28.9 - -17.9]	-4.8	[-8.9 - -0.8]	-32.8	[-39.8 - -26.6]
RR HG-average	0.22	[0.13 - 0.3]	0.25	[0.02 - Inf]	0.23	[0.14 - 0.32]
RR HG-max	Inf	[Inf - Inf]	Inf	[0.01 - Inf]	Inf	[Inf - Inf]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Burden (%)	-1.2	[-1.5 - -0.9]	2.3	[1.9 - 2.8]	7.8	[6.0 - 9.9]
Prevalence (%)	25.0	[20.1 - 29.8]	34.4	[28.3 - 40.5]	46.8	[38.5 - 55.1]
Coverage (%)	31.1	[24.8 - 37.5]	56.6	[48.3 - 64.9]	17.5	[13.2 - 21.8]
Relative benefit (%)	118.4	[77.7 - 180.4]	374.3	[-4986.7 - 5868.4]	175.7	[129.4 - 234.6]
Quality (%)	-122.0	[-167 - -84.1]	-49.7	[-82.2 - -21.3]	900.4	[661.4 - 1205.3]
Relative - effective coverage (%)	47.5	[29 - 86.7]	187.4	[-1478.2 - 2065.6]	-71.8	[-430.6 - -25.8]
Absolute - effective coverage (%)	-1.5	[-2.0 - -1.0]	-3.3	[-5.2 - -1.4]	-2.3	[-3.0 - -1.7]
Effective coverage (%)	-37.6	[-51.7 - -25.9]	-28.8	[-46.9 - -12.3]	145.4	[104.2 - 199.1]
RR HG-average	Inf	[4.12 - Inf]	Inf	[0.02 - Inf]	Inf	[Inf - Inf]
RR HG-max	0.51	[0.41 - 0.58]	0.11	[0.01 - Inf]	1.24	[1.18 - 1.34]

The coverage of the treatment for hypertension, presented in Table 5.7, varied from 12.7% in Ghana to 56.6% in the Russian Federation.

Almost all values of effective coverage calculated using the methodological proposal of this thesis were negative. In part, this is because, in all countries, the health gain estimated for people receiving treatment was below zero (see Table 5.8). Figure 5.4 shows the distribution of health gains for the study countries, and it can be observed that health gains tend to be displaced towards negative numbers. Negative health gains mean that people receiving treatment for hypertension have a higher disability (or lower health) than people without treatment, regardless of the variables used to adjust the models (see Table S5.16 from the supplementary material for results of the regression models and Figure S5.20 for histograms of health gains separately for each country).

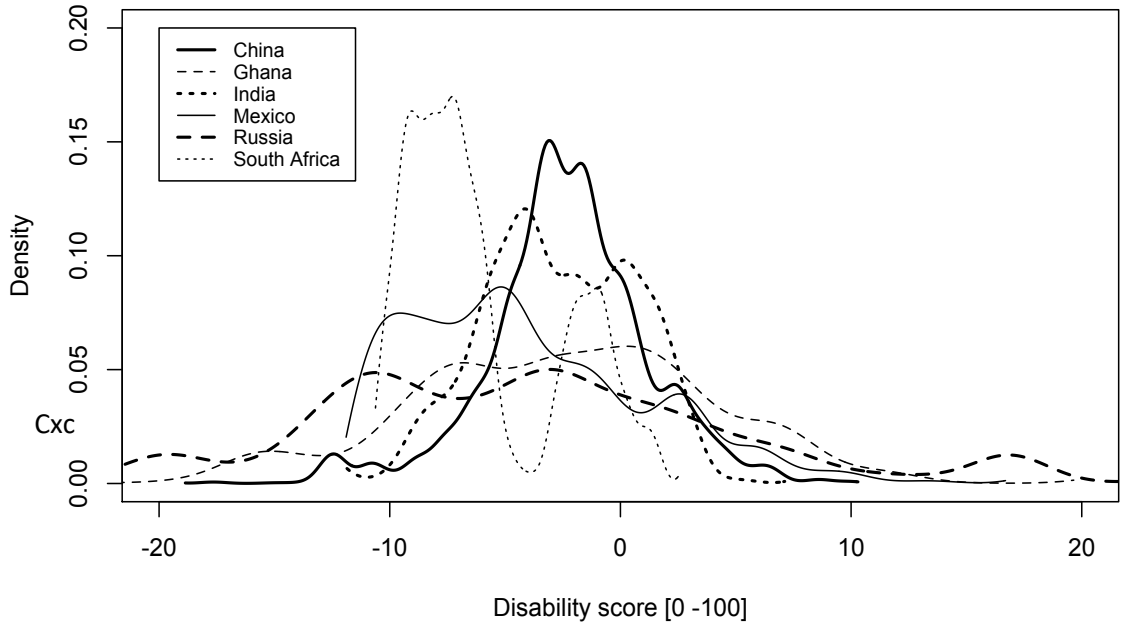
Table 5.8. Disability attributable to depressive disorder according to the analysis of the burden of disability, and results from the effective coverage analysis of healthcare for hypertension expressed in person-years. SAGE study, Wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
n weighted	270,697,618	-	11,052,687	-	279,003,445	-
Disability Attributable	502,366	[467214 - 536904]	3,977	[3624 - 4336]	831,647	[779157 - 885505]
DA_{G1}	138,055	[-12368 - 288477]	-7,027	[-23944 - 9891]	479,520	[379352 - 579688]
DA'_{G2}	141,801	[81241 - 202360]	2,378	[-3387 - 8143]	81,027	[43540 - 118514]
HG_{G2}	-502,614	[-615711 - -389518]	-12,305	[-22583 - -2028]	-266,563	[-316012 - -217114]
HG'_{G1}	1,617,539	[1487785 - 1747293]	223,790	[202454 - 245125]	634,460	[572225 - 696696]
HG'_{G2}	544,954	[482634 - 607273]	32,686	[24920 - 40452]	179,277	[156169 - 202384]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
n weighted	53,571	-	83,442,881	-	21,452,752	-
Disability Attributable	-138	[-160 - -115]	471,963	[399486 - 545125]	362,776	[291267 - 432634]
DA_{G1}	-215	[-361 - -69]	-192,105	[-538596 - 154386]	202,098	[65669 - 338527]
DA'_{G2}	-143	[-183 - -103]	-107,272	[-341613 - 127069]	-60,151	[-70887 - -49414]
HG_{G2}	-170	[-220 - -119]	-639,043	[-1000014 - -278073]	-105,486	[-129731 - -81241]
HG'_{G1}	312	[242 - 382]	938,385	[674985 - 1201786]	-61,118	[-76255 - -45982]
HG'_{G2}	139	[116 - 161]	1,289,475	[991651 - 1587298]	-11,721	[-13913 - -9530]

As I commented in the case of depression, negative health gains imply that the effective coverage calculation and all other indicators associated with it will also have negative values. However, when negative health values are constrained according to the sensitivity analysis (see Table 5.9), at least in China and India, positive values in all the indicators of effective coverage are obtained. In both countries, the raw coverage was around 20%, while the quality indicator was close to 20%, and the effective coverage was about 5%. The difference in the relative benefit indicator between China and India is due to the size of the disability associated with hypertension. Moreover, since the prevalence of hypertension is noticeably different between both countries, the a-EC difference is also large. Extreme values in the RR HG-maximum were produced due to the maximum health gain overpassing the disability attributable to hypertension (see Table 5.8). Again, this is a counterintuitive result, which is not sensitive enough to constrain the negative values of health gains.

Figure 5.4. Observed distribution of health gain attributable to the healthcare intervention addressed to hypertension, by country. SAGE study, Wave 1.



5.5.2. Negative values in HG and a negative attributable disability in people without treatment

In the case of Ghana, in addition to a negative health gain associated with the treatment of hypertension, we observed a negative disability attributable to the disease in people without treatment (see Table 5.8). Restricting negative values of health gain (sensitivity analysis), as in the case of China and India, turns the indicators of effective coverage into positive numbers. Quality is re-estimated at 21.9%, and the effective coverage reached 2.8% (see Table 5.9). This is consistent with the product of coverage (12.7%) and quality.

However, the r-EC remains negative, since it is very dependent on the disability attributable to the disease in people without treatment. Additionally, it highlights a relative benefit higher than 100%, which again is an unexpected result.

Table 5.9. Results of effective coverage of healthcare for hypertension, restricting negative values of health gain (sensitivity analysis). SAGE study, Wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	78.3	[49.5 - 141.2]	175.8	[-2620.6 - 2909.1]	46.2	[23.2 - 94.9]
Quality (%)	20.5	[14.8 - 26.6]	21.9	[10.2 - 36.8]	20.9	[11.6 - 30.8]
Relative - effective coverage (%)	39.7	[22.8 - 97.6]	-55.7	[-1063.9 - 907.6]	6.7	[3.7 - 10.1]
Absolute - effective coverage (%)	0.20	[0.1 - 0.2]	0.29	[0.1 - 0.5]	0.05	[0 - 0.1]
Effective coverage (%)	5.1	[3.8 - 6.5]	2.8	[1.3 - 4.3]	4.6	[2.6 - 6.7]
RR HG-average	4.79	[1.94 - Inf]	Inf	[0.06 - Inf]	1.85	[1.31 - 17.95]
RR HG-max	Inf	[Inf - Inf]	Inf	[0.01 - Inf]	Inf	[Inf - Inf]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	-23.4	[-35.5 - -15.6]	-187.3	[-3140.4 - 2470.2]	-1.3	[-2.3 - -0.3]
Quality (%)	24.1	[16.9 - 32.7]	25.0	[4.9 - 48.1]	99.9	[20 - 382.5]
Relative - effective coverage (%)	-9.4	[-17.1 - -5.8]	-90.5	[-996.2 - 790.1]	0.5	[0 - 3.3]
Absolute - effective coverage (%)	0.29	[0.2 - 0.4]	1.65	[0.3 - 3]	0.02	[0 - 0]
Effective coverage (%)	7.4	[5.2 - 10.2]	14.5	[2.8 - 27.2]	43.8	[10.7 - 93]
RR HG-average	0.81	[0.74 - 0.86]	0.36	[0.03 - Inf]	0.99	[0.98 - 1]
RR HG-max	0.51	[0.41 - 0.58]	0.11	[0.01 - Inf]	0.99	[0.98 - 1]

5.5.3. Negative values in HG and a negative attributable disability in people with and without treatment

Results in Table 5.8 shows similar patterns for Mexico and the Russian Federation. The HG numbers are negative, and so too the disability attributable to hypertension among people with and without treatment. All these results are unexpected and hardly plausible. However, the sensitivity analysis – where negative values of HG are assumed to be zero (see Table 5.9) – leads to results for effective coverage that seem reasonable and consistent with the product of the coverage and the indicator of quality. However, this is only in appearance, because they result from operations using non-plausible values (i.e. the disease is associated with less disability, and the treatment is associated with higher disability).

5.5.4. Negative values in the maximum HG

South Africa was the only country where almost all estimates of individual health gain were negative, including the maximum health gain. In this case, when negative values were replaced with zero (sensitivity analysis), the variability of the health gains decreased greatly,

which explains a quality close to 100% (see Table 5.9). The disability attributable to the disease in those people with treatment was also negative.

In summary, in the case of hypertension, where the disability associated with the disease was generally small, the procedure used to estimate effective coverage did not work properly. Restricting negative values of health gain produced more reasonable results. However, other indicators, such as the relative benefit and the r-EC, suggest that these operations are made using meaningless values. Results assuming the effective coverage as an individual attribute are presented in the supplementary Table S5.17, but the results remained inconsistent (second sensitivity analysis).

5.6. Effective coverage of the healthcare addressing to osteoarthritis

Osteoarthritis is usually reported in the affected joint and prevalence levels vary notably according to the method used to detect cases. In adults, using questionnaires of symptoms, the prevalence for osteoarthritis of the knee has been estimated between 5.4% and 6.3%, while for the hip joint it is between 0.8% and 1.6%. Hand's osteoarthritis has been calculated at around 2% [22]. According to estimates from the Global Burden of Disease project, the prevalence levels for knee and hip osteoarthritis are about 3.6% and 0.85%, respectively, for the entire population [23]. In the SAGE study, the prevalence of any osteoarthritis condition varied between 6.2% in Mexico and 13.5% in the Russian Federation (see Table 5.10).

Mexico showed the smallest burden of disability attributable to osteoarthritis (1.9%), while Ghana presented the largest burden (5.2%); see Table 5.10. The high burden observed in Ghana can be explained by the fact that the disability associated with osteoarthritis was highest among the countries studied (see Table 5.3). Compared with other estimates of the burden from the IHME, based on years lived with disability, results are highly similar for China, India, Mexico and the Russian Federation [IHME: https://vizhub.healthdata.org/gbd-compare/ consulted in February 2021 for the year 2007]. However, for Ghana and South Africa, the IHME estimated smaller burdens.

Table 5.10. Burden of disability, prevalence, coverage of the treatment and results of effective coverage of the healthcare for osteoarthritis, by country. SAGE study, Wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Burden (%)	2.7	[2.3 - 3.1]	5.2	[4.4 - 6.1]	2.1	[1.9 - 2.3]
Prevalence (%)	8.4	[7.2 - 9.6]	9.8	[8.2 - 11.4]	8.7	[7.9 - 9.5]
Coverage (%)	53.5	[46 - 61]	24.9	[17.8 - 32.1]	43.2	[38.6 - 47.8]
Relative benefit (%)	-98.6	[-182.4 - -54.5]	-18.8	[-79.8 - 16.8]	-26.8	[-40.6 - -14.2]
Quality (%)	-129.5	[-188.5 - -77.8]	-25.2	[-80.8 - 21]	-37.8	[-56.5 - -20.2]
Relative - effective coverage (%)	-49.6	[-76.9 - -28.9]	-3.4	[-10 - 2.8]	-12.0	[-17.9 - -6.4]
Absolute - effective coverage (%)	-0.9	[-1.3 - -0.5]	-0.2	[-0.4 - 0.1]	-0.2	[-0.3 - -0.1]
Effective coverage (%)	-69.5	[-99.4 - -42]	-6.3	[-18.3 - 5.2]	-16.5	[-24.3 - -8.9]
RR HG-average	0.51	[0.36 - 0.65]	0.84	[0.55 - 1.19]	0.79	[0.72 - 0.87]
RR HG-max	4.15	[2.09 - Inf]	3.84	[1.56 - Inf]	3.40	[2.4 - 7.6]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Burden (%)	1.9	[1.3 - 2.6]	3.1	[2.2 - 4.1]	4.3	[3 - 5.8]
Prevalence (%)	6.2	[4.1 - 8.4]	13.5	[9.5 - 17.6]	9.9	[6.8 - 13]
Coverage (%)	47.6	[30.9 - 64.2]	72.2	[62.8 - 81.6]	60.3	[44.3 - 76.4]
Relative benefit (%)	-25.3	[-616.8 - 447.9]	18.9	[-73 - 158]	-81.7	[-239.9 - 6.4]
Quality (%)	-7.0	[-81.7 - 65.1]	8.8	[-23.3 - 43.5]	-34.0	[-76.3 - 2.6]
Relative - effective coverage (%)	-8.9	[-114.7 - 91.1]	14.7	[-45.4 - 91]	-45.6	[-120.1 - 3.5]
Absolute - effective coverage (%)	-0.2	[-1.6 - 1.3]	0.5	[-1.2 - 2.2]	-1.4	[-3 - 0.1]
Effective coverage (%)	-3.2	[-33.5 - 27.3]	6.3	[-16.2 - 29.8]	-20.7	[-47.8 - 1.5]
RR HG-average	0.81	[0.13 - Inf]	1.23	[0.58 - Inf]	0.54	[0.29 - 1.05]
RR HG-max	Inf	[1.38 - Inf]	Inf	[8.95 - Inf]	Inf	[Inf - Inf]

The coverage of the treatment for osteoarthritis ranged between 24.9% observed in Ghana and 72.2% in the Russian Federation (see Table 5.10). The level of coverage, overall, is more extensive than those observed for depression and hypertension. However, the item in the questionnaire that asked about the treatment was formulated broadly, which may include a spectrum of interventions, from orthoses to the self-prescription of analgesics.

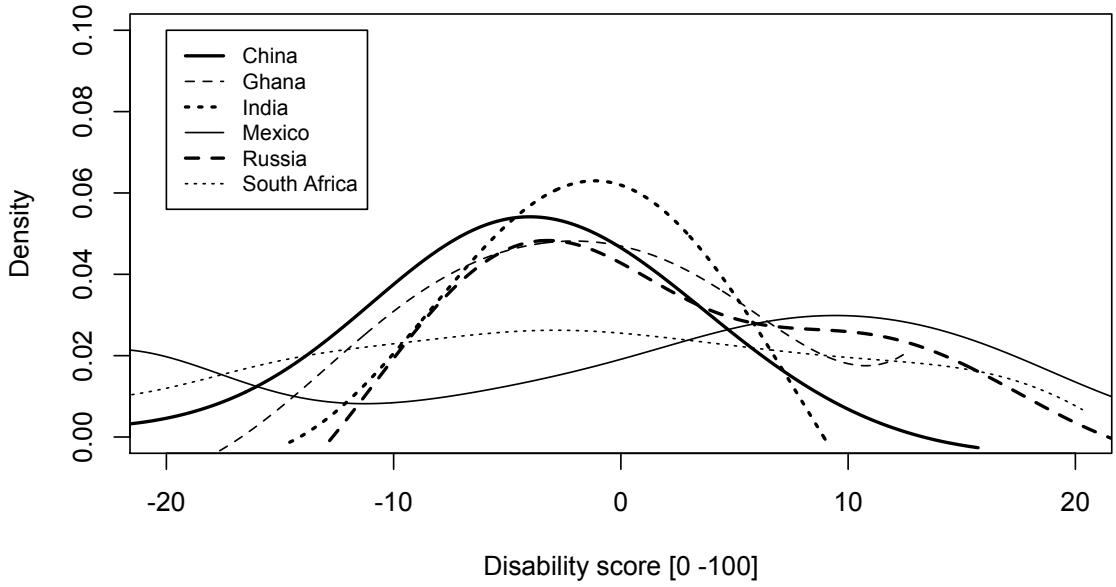
Table 5.11. Disability attributable to osteoarthritis according to the burden analysis, and results of effective coverage of healthcare for osteoarthritis, expressed in person-years. SAGE study, Wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
n weighted	270,697,618	-	11,052,687	-	279,003,445	-
Disability Attributable	1,562,236	[1337171 - 1789517]	128,162	[108127 - 148186]	1,596,585	[1477174 - 1720676]
DA_{G1}	509,380	[358440 - 660319]	96,601	[68021 - 125181]	774,279	[662256 - 886302]
DA'_{G2}	517,922	[306399 - 729444]	20,611	[7509 - 33713]	627,438	[531721 - 723154]
HG_{G2}	-507,988	[-707267 - -308710]	-3,896	[-11012 - 3220]	-167,722	[-245212 - -90232]
HG'_{G1}	339,347	[268672 - 410022]	46,220	[36663 - 55778]	573,389	[518383 - 628394]
HG'_{G2}	392,936	[328918 - 456954]	15,543	[9641 - 21444]	443,513	[390842 - 496185]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
n weighted	53,571	-	83,442,881	-	21,452,752	-
Disability Attributable	221	[153 - 289]	636,597	[452710 - 822744]	201,008	[144423 - 255971]
DA_{G1}	144	[25 - 264]	157,855	[104773 - 210936]	61,591	[11453 - 111728]
DA'_{G2}	67	[1 - 134]	514,056	[86512 - 941600]	76,640	[34994 - 118286]
HG_{G2}	-17	[-184 - 150]	102,526	[-250218 - 455271]	-63,133	[-132624 - 6358]
HG'_{G1}	320	[185 - 455]	438,940	[348366 - 529514]	119,679	[47827 - 191532]
HG'_{G2}	266	[128 - 404]	1,138,208	[686007 - 1590410]	183,763	[138443 - 229083]

Results of effective coverage and their related indicators also are presented in Table 5.10. Results of the parameters expressed in person-years of disability are shown in Table 5.11. Most of the health gains were estimated in negative ranges (see Table 5.11). The distribution of health gain by country is presented in Figure 5.5, where it can be observed that this parameter covers an ample area in the scale of disability (see Figure S5.21 in the supplement for separate histograms for each country). This variability can represent a difference in the effects of healthcare interventions in each individual, or a difference in the quality of the treatment (see chapter four for further explanation).

Figure 5.5. Observed distribution of health gain attributable to healthcare interventions addressed to osteoarthritis, by country. SAGE study, Wave 1.



As I discussed for depression and hypertension, a consequence of negative health gains is that all the indicators of effective coverage also show negative values in Table 5.10, which cannot be considered plausible. The exception was the Russian Federation, which showed a positive health gain in Table 5.11, and consequently showed the quality of healthcare above zero (8.8%) and effective coverage of 6.3%.

The results from the sensitivity analysis are presented in Table 5.12, where negative values of health gains were restricted. The resulting values for the indicators seem more plausible, except for the relative benefit in Mexico, where the central estimation is higher than the unit. The quality of the healthcare intervention for osteoarthritis varied between 36.5% in Mexico – although it could be overestimated, since it shows relative benefits higher than 100% – and 22.9% in India.

The highest effective coverage is still observed in the Russian Federation (19.2%), while the smallest is observed in Mexico (1.5%), despite a possible overestimation of the relative benefit.

Table 5.12. Results of effective coverage of the healthcare for osteoarthritis, restricting negative values of health gain (sensitivity analysis). SAGE study, Wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	18.4	[12.2 - 32.2]	25.9	[1.7 – 83.0]	16.2	[11.3 - 21.9]
Quality (%)	24.3	[18.8 - 31.1]	34.1	[2.7 - 77.1]	22.9	[16.1 - 30.3]
Relative - effective coverage (%)	9.2	[6.7 - 13.0]	4.5	[0.3 - 9.4]	7.3	[5.1 - 9.6]
Absolute - effective coverage (%)	0.17	[0.1 - 0.2]	0.22	[0 .0- 0.4]	0.13	[0.1 - 0.2]
Effective coverage (%)	13.0	[10.2 - 16.3]	8.6	[0.7 - 16.9]	10.0	[7.1 - 13.0]
RR HG-average	1.22	[1.14 - 1.47]	1.35	[1.01 - 6.02]	1.19	[1.13 - 1.29]
RR HG-max	4.11	[2.01 - Inf]	4.25	[1.64 - Inf]	3.42	[2.39 - 6.98]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	136.6	[-31.3 - 928.3]	58.3	[-18.7 - 314.4]	58.6	[22 - 142.7]
Quality (%)	36.5	[7.9 - 89.7]	26.6	[-4.1 - 66.5]	24.4	[9.5 - 41.3]
Relative - effective coverage (%)	45.6	[9.6 - 144.8]	45.4	[-8.3 - 163.5]	32.5	[12.1 - 70.5]
Absolute - effective coverage (%)	0.83	[0.2 - 1.5]	1.49	[-0.2 - 3.2]	0.97	[0.4 - 1.6]
Effective coverage (%)	1.5	[3.6 - 33]	19.2	[-3.0 - 44.8]	14.8	[5.8 - 26.2]
RR HG-average	Inf	[0.99 - Inf]	2.39	[0.89 - Inf]	2.38	[1.27 - Inf]
RR HG-max	Inf	[7.51 - Inf]	Inf	[15.72 - Inf]	Inf	[Inf - Inf]

Results assuming effective coverage as an attribute of individuals (second sensitivity analysis) are presented in the supplementary Table S5.19, which are in overall agreement with results from Table 5.10.

Again, in summary, with the proposed approach negative estimates of health gains remain a problem for the calculation of effective coverage and other indicators. As in the case of depression, when negative values are restricted, the results seem more reasonable. However, the justification of this constraint is weak (see discussion in chapters four and six).

5.7. How to use effective coverage indicators in the assessment of healthcare performance across countries and healthcare interventions

This chapter has explored the performance of a new approach to estimate effective coverage across different healthcare interventions and countries. I have shown that the main problem of the procedure is the estimation of negative health gains associated with the treatment. I will further develop this limitation in chapter six.

Despite this limitation, one of the strengths of the procedure is the number of new indicators proposed that are related to the concept of effective coverage. These indicators can potentially provide valuable information about different aspects of the performance of healthcare services. This section will offer a comprehensive account of how to use these indicators in that regard. The focus will be on comparing performance across countries and healthcare interventions.

For this, I will use the results already shown for depression and osteoarthritis, according to the sensitivity analysis, where negative health gains were assumed to equal zero.

5.7.1. Effectiveness of healthcare

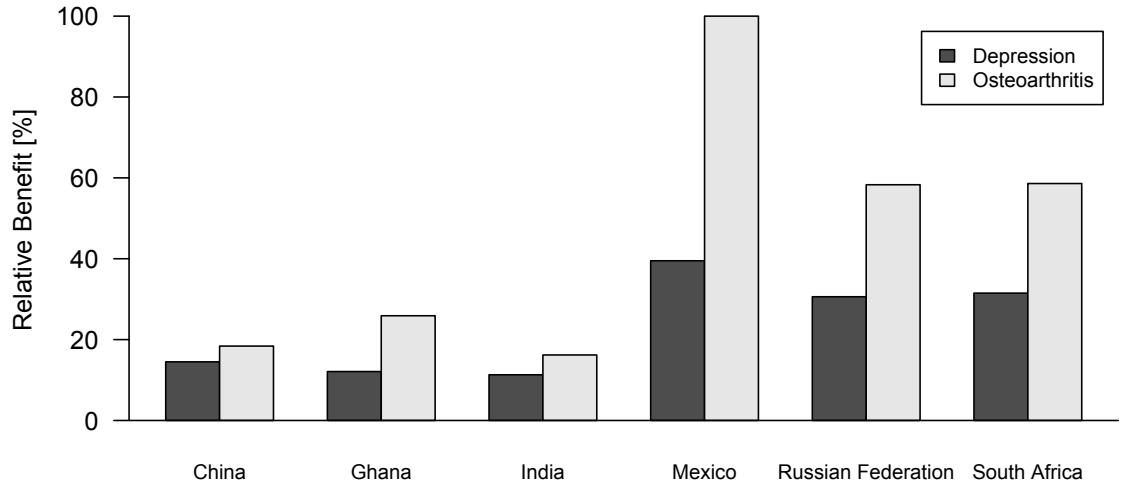
Two indicators can be used to measure the overall effectiveness of healthcare interventions across diseases and countries: relative benefit and RR HG-maximum.

The first, relative benefit, can be expressed as an RR HG-average (see chapter four). Figure 5.6 visually compares the relative benefit for depression and osteoarthritis by country, using the results with restricted negative health gains (data extracted from Tables 5.6 and 5.12).

It can be clearly observed that, in people who received the healthcare intervention, there was a higher relative benefit for osteoarthritis than for depression, in terms of avoided disability. It is important to remember that the magnitude of disability attributable to each health condition is specific for each country. In other words, the disability weights used for each disease are country specific.

Additionally, a general trend of higher relative benefit can be seen in Mexico, the Russian Federation, and South Africa than in China, Ghana and India. However, the uncertainty intervals are large and preclude making clear distinctions between countries.

Figure 5.6. Comparison of the effectiveness of healthcare interventions for depression and osteoarthritis by country, using the relative benefit indicator from the effective coverage analysis and restricting negative health gains. SAGE study, Wave 1.

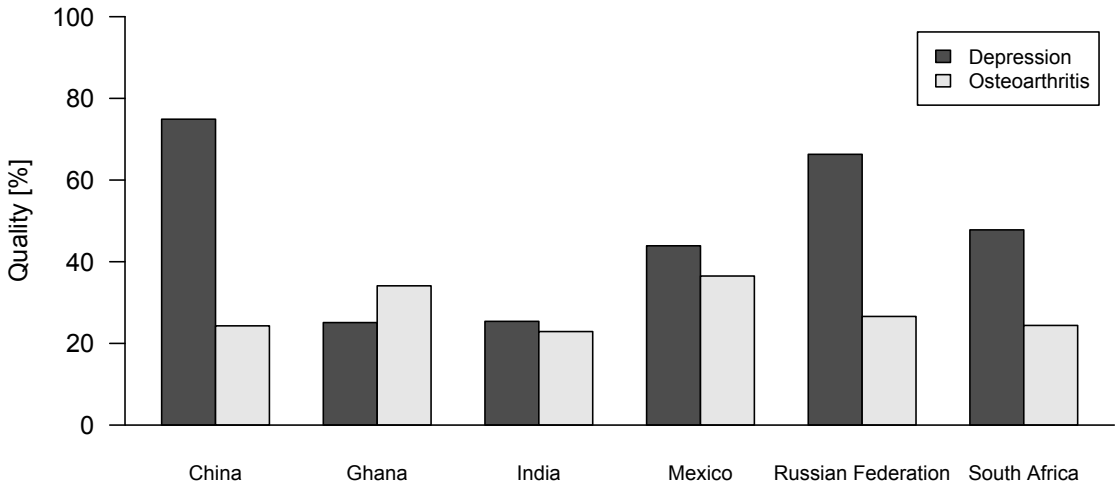


Another indicator that also measures the effectiveness of the healthcare intervention is the RR HG-maximum, which is equivalent to relative benefit, but assumes that all the population obtained the highest benefit from the healthcare intervention. As mentioned, it represents the ceiling of the benefit that a country can offer to their population. This indicator can be related to the level of health technology development for each country.

5.7.2. Quality of healthcare

The proposal also offers an indicator about the quality of the healthcare services (see Figure 5.7). Contrary to the case of effectiveness, we can see a trend of higher quality healthcare interventions for depression than for osteoarthritis. Also, it is possible to observe that China, a country where the effectiveness of the healthcare intervention for depression was relatively small, has the highest values for quality.

Figure 5.7. Comparison of the quality of healthcare interventions for depression and osteoarthritis by country, restricting negative health gains. SAGE study, Wave 1.



The quality of healthcare interventions addressing osteoarthritis showed a more homogeneous level throughout the countries. It is worth remembering that the quality is calculated using the maximum health gain from each country; therefore, the quality indicators are not strictly comparable. This issue can be resolved using a unique model including all countries and calculating a single maximum health gain (for instance, through a mixed-effect model). This issue was not explored because the chapter aimed to assess only the performance of the overall procedure within different countries, using their own parameters.

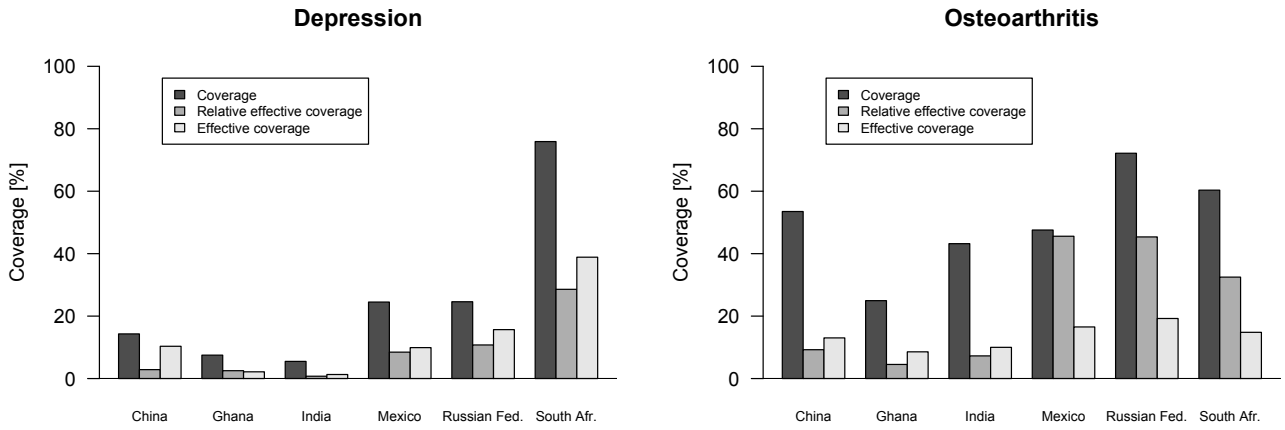
5.7.3. Combing coverage with effectiveness and quality

Some indicators combine coverage with effectiveness (i.e. the r-EC), while others coverage with quality (i.e. effective coverage); see Figure 5.8.

In the case of depression, it can be appreciated that raw coverages are small, except for South Africa. Moreover, the gap between the raw coverage and the r-EC and effective coverage is

high. South Africa shows the best raw coverage and the best performance in terms of the central estimation for r-ECe and effective coverage.

Figure 5.8. Comparison between raw coverage, relative effective coverage, and effective coverage of healthcare interventions for depression and osteoarthritis by country, restricting negative health gains. SAGE study, Wave 1.



Russian Fed: Russian Federation; South Afr.: South Africa

In the case of osteoarthritis, overall raw coverage is higher than for depression. In some countries, such as Mexico, the Russian Federation and South Africa, the gaps between raw coverage and r-EG are small, at least in comparison with China, Ghana, and India. However, in terms of effective coverage, in general, all countries show central estimates below 20%.

The relative effective coverage and the effective coverage seem to be complementary indicators. While the first one depends on the coverage and the average effectiveness of the healthcare intervention, the second indicator depends on the coverage and the variability of the effectiveness. To understand better the relationship between both, Figure S5.21 from the supplementary material shows simulated results for relative benefit, quality, relative effective coverage and effective coverage, assuming different distributions of health gains.

5.7.4. The most comprehensive indicator

Finally, the a-EC provides the avoided disability attributable to the healthcare intervention, using as a denominator the total burden of disability from each country. This indicator allows a direct comparison between different healthcare interventions and countries. This indicator brings into a single expression the disability attributable to disease (i.e. the disability weight), the prevalence of the disease, the coverage, and the effectiveness attributable to the healthcare intervention.

Figure 5.9. Comparison of the absolute effective coverage of healthcare interventions for depression and osteoarthritis by country, restricting negative health gains. SAGE study, Wave 1.

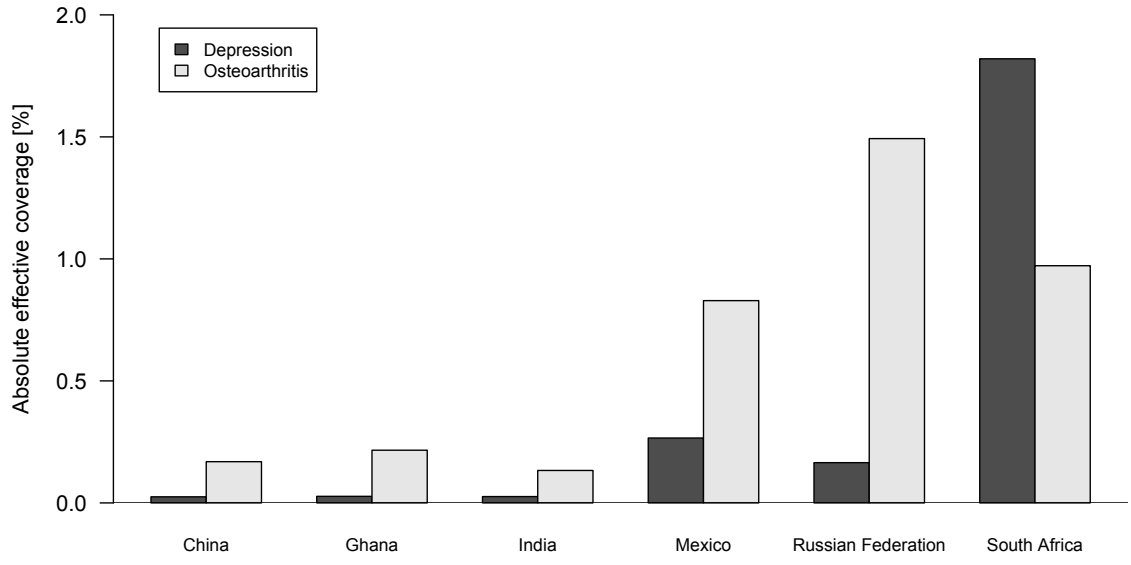


Figure 5.9 shows the a-EC for depression and osteoarthritis for each country in the study. The healthcare intervention for depression in South Africa averted the highest fraction of the national burden of disability. It is almost twice the burden avoided by the osteoarthritis healthcare intervention in that country. This is relevant because the raw coverage for both healthcare interventions is not so different (75.9% versus 60.3% for depressive disorder and osteoarthritis, respectively).

We also can observe that, in the Russian Federation, the disability avoided by the healthcare intervention for osteoarthritis is nine times higher than the disability averted for depressive

disorder, even though the raw coverage is only three times higher (72.2% versus 24.6% for osteoarthritis and depressive disorder, respectively).

In China, Ghana and India, overall, the disability avoided by the healthcare interventions is smaller than in the other three countries in the study. In China, these results were obtained even when the country showed a high raw coverage for depressive disorder.

The a-EC can also be combined with the budgets addressed to each healthcare intervention, informing the cost of each disability unit averted or the value of money in terms of disability. Moreover, as was previously pointed out, it is the indicator that produces the least uncertainty.

5.8. Conclusions of the chapter

In this chapter, I have shown that the procedure proposed to measure effective coverage is limited in its ability to generate reliable estimates of the health gains associated with healthcare interventions, especially in diseases associated with a small level of disability. In addition, some results tend to be estimated with a considerable level of uncertainty.

On the other hand, I have shown that several indicators emerge from the main approach to measure effective coverage. From a theoretical point of view, these indicators offer an interesting method of measuring different aspects of the health services: effectiveness, quality, and the result of their combination with coverage. Moreover, one of these indicators (i.e. a-EC) also includes information about prevalence, expressed by the same unit metric across different diseases. As a consequence, it might be especially suitable for decision making.

In addition, I have shown that the metric of effective coverage is consistent with the metric for calculating disability attributable to diseases (chapter three), and that both are based on a strong conceptual framework for evaluating health states (see chapter two).

In the next chapter, I will discuss the overall limitations of the procedure to calculate effective coverage, with a focus on the challenges for future research. Furthermore, I will assess how well this proposal of effective coverage answers the aims of the thesis.

References of chapter five

1. Shengelia B, Tandon A, B. Adams O, Murray CJL: **Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy**. *Social Science & Medicine* 2005, **61**:97-109.
2. Ng M, Fullman N, Dieleman JL, Flaxman AD, Murray CJL, Lim SS: **Effective Coverage: A Metric for Monitoring Universal Health Coverage**. *PLOS Medicine* 2015, **11**(9).
3. World Health Organization: **The World Health Report. Health Systems Financing. The path to universal coverage**. Switzerland: World Health Organization; 2010.
4. Fullman N, Barber RM, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, Abdulkader RS, Abdulle AM, Abera SF *et al*: **Measuring progress and projecting attainment on the basis of past trends of the health-related Sustainable Development Goals in 188 countries: an analysis from the Global Burden of Disease Study 2016**. *The Lancet* 2017, **390**(10100):1423-1459.
5. World Health Organization and the International Bank for Reconstruction and Development / The World Bank: **Tracking universal health coverage: 2017 global monitoring report**. Switzerland.; 2017.
6. Kowal P, Chatterji S, Naidoo N, Biritwum R, Fan W, Lopez Ridaura R, Maximova T, Arokiasamy P, Phaswana-Mafuya N, Williams S *et al*: **Data resource profile: the World Health Organization Study on global AGEing and adult health (SAGE)**. *Int J Epidemiol* 2012, **41**(6):1639-1649.
7. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, Angermeyer MC, Bernert S, de Girolamo G, Morosini P *et al*: **Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys**. *JAMA* 2004, **291**(21):2581-2590.
8. Association AP: **Diagnostic and Statistical Manual of Mental Disorders**. Washington, DC: American Psychiatric Association; 1994.
9. Arokiasamy P, Uttamacharya U, Jain K, Biritwum RB, Yawson AE, Wu F, Guo Y, Maximova T, Espinoza BM, Rodriguez AS *et al*: **The impact of multimorbidity on adult physical and mental health in low- and middle-income countries: what does the study on global ageing and adult health (SAGE) reveal?** *BMC Med* 2015, **13**:178.
10. Wu F, Guo Y, Chatterji S, Zheng Y, Naidoo N, Jiang Y, Biritwum R, Yawson A, Minicuci N, Salinas-Rodriguez A *et al*: **Common risk factors for chronic non-communicable diseases among older adults in China, Ghana, Mexico, India, Russia and South Africa: the study on global AGEing and adult health (SAGE) wave 1**. *BMC Public Health* 2015, **15**:88.
11. Üstün TB, Chatterji S, Mechbal A, Murray CJL, Groups WC: **The World Health Surveys**. In: *Health Systems Performance Assessment Debates, Methods and Empiricsm.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 797 - 808.

12. Humphreys LG, Montanelli Jr RG: **An Investigation of the Parallel Analysis Criterion for Determining the Number of Common Factors**. *Multivariate Behavioral Research* 1975, **10**(2):193-205.
13. Salomon JA, Mathers CD, Chatterji S, Sadana R, Üstün TB, Murray CJL: **Quantifying Individual Levels of Heath: Definitions, Concepts, and MEasurement Issues**. In: *Health Systems Performance Assessment Debates, Methods and Empiricism.* edn. Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003: 301-318.
14. Satorra A, Bentler PM: **Ensuring Positiveness of the Scaled Difference Chi-square Test Statistic**. *Psychometrika* 2010, **75**(2):243-248.
15. Hoyle RH: **Handbook of structural equation modeling**: Guilford Publications; 2012.
16. Cronbach LJ: **Coefficient alpha and the internal structure of tests**. *Psychometrika* 1951, **16**(3):297-334.
17. Greenland S, Pearce N: **Statistical foundations for model-based adjustments**. *Annu Rev Public Health* 2015, **36**:89-108.
18. Rai D, Zitko P, Jones K, Lynch J, Araya R: **Country- and individual-level socioeconomic determinants of depression: multilevel cross-national comparison**. *Br J Psychiatry* 2013, **202**(3):195-203.
19. Araya R, Zitko P, Markkula N: **The Impact of Universal Health Care Programmes on Improving 'Realized Access' to Care for Depression in Chile**. *Adm Policy Ment Health* 2018, **45**(5):790-799.
20. Sarki AM, Nduka CU, Stranges S, Kandala NB, Uthman OA: **Prevalence of Hypertension in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis**. *Medicine (Baltimore)* 2015, **94**(50):e1959.
21. Salomon JA, Haagsma JA, Davis A, de Noordhout CM, Polinder S, Havelaar AH, Cassini A, Devleesschauwer B, Kretzschmar M, Speybroeck N *et al*: **Disability weights for the Global Burden of Disease 2013 study**. *The Lancet Global health* 2015, **3**(11):e712-723.
22. Pereira D, Peleteiro B, Araujo J, Branco J, Santos RA, Ramos E: **The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review**. *Osteoarthritis Cartilage* 2011, **19**(11):1270-1285.
23. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, Bridgett L, Williams S, Guillemin F, Hill CL *et al*: **The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study**. *Ann Rheum Dis* 2014, **73**(7):1323-1330.

Chapter 6: Overall Discussion

6.1. Introduction

Each of the previous chapters has included a section of discussion which has addressed the methodological aspects of the different steps taken to develop a proposal for measuring effective coverage. In this chapter I will provide a more general discussion, aiming to present an overview of the strengths, limitations, and future challenges of the proposal, framed within the context that originally motivated this thesis.

In the second section, I will present a summary of the results obtained in each of the previous chapters. Then, in the third section, I will discuss how well these results answer the main research question of the thesis. Next, across the fourth and fifth sections, I will present the strengths of the proposed metric, as well as the future challenges. Finally, in the sixth section, I will discuss the convenience of using effective coverage as an indicator for tracking progress in Universal Health Coverage (UHC). The chapter ends with an overall conclusion.

6.2. A brief account of the previous chapters

The previous chapters have progressively described and tested a new proposal for measuring the effective coverage of healthcare interventions, which might be suitable for monitoring UHC.

Chapter one introduced the concept of effective coverage according to its two leading developers, Tanahashi and Shengelia et al. [1, 2]. I reviewed how well this concept fits with its empirical use in the scientific literature, as well as with the measurement of UHC in the context of the Sustainable Development Goals (SDG) [3]. The main conclusion of the first chapter was that effective coverage is not usually measured as intended. The main reason is the lack of information about the quality of health services. In this context, quality was defined as the fraction of the maximum potential health gain achieved by people through a healthcare intervention. In addition, effective coverage has been mostly used to assess those healthcare interventions addressed to infectious diseases and maternal-child and nutritional conditions in low-and-middle-income countries.

Chapter two introduced the concept of health. The ideas of health status, health states, disability, and health utilities were discussed as potential outcomes for measuring effective coverage. Chapter two argued that, through a regression model applied to a latent variable of disability – such as health status, or health utilities – it is possible to estimate disability weights (or health-utility weights) for different health states associated with a disease, adjusting by comorbidities and other confounders.

In chapter three I used the same procedure presented in chapter two for calculating disability weights, to propose a new way of calculating the burden of disability (or the loss of health-state utilities) associated with diseases. That procedure was practical, straightforward and original, based on the concept of attributable fractions and applied to continuous variables. Using this proposal to measure the burden of disability could also have several advantages when compared with the standard methodological alternative (i.e. years lived with disability). Two main approaches were then suggested: the population average-level and the individual-level.

In chapter four the procedure to calculate the burden of disability was explored, in order to estimate the fraction of avoidable disability attributable to a disease, which can be avoided via a healthcare intervention. I discussed the complementary results obtained through the population average-level and individual-level approaches. Also, various indicators associated with effective coverage were raised: health benefit, quality, relative effective coverage (r-EC), absolute effective coverage (a-EC), and the effective coverage itself. A first assessment of the proposed approach was implemented using data from one country and one healthcare intervention as a case study. The main assumptions of the proposal were identified and discussed.

Finally, in chapter five, the methodological proposal to measure effective coverage was applied to data from different countries and several healthcare interventions. A comprehensive appraisal of the performance of the procedure was carried out. Furthermore, the best way to employ and interpret each indicator associated with effective coverage was discussed.

6.3. Is the research question answered?

The research question of this thesis concerns the feasibility of creating a new approach for measuring effective coverage, in an effort to overcome the limitations of Shengelia et al.´s framework.

The main challenge that the research question places on this new approach was how to include a component of quality based on the concept of potential health gain attributable to a healthcare intervention.

The proposal to measure effective coverage presented in this thesis complies with the inclusion of a quality component based on the concept of health gain. Furthermore, throughout chapters two to four, it was argued that the way health gains are included in the metric of effective coverage and the conceptual framework of health state valuation were both consistent. Also, it was emphasised that, through the proposed procedure, the outcome of health gain could be easily changed to other alternatives more suitable for decision-makers – for instance, health-state utilities.

However, there are three important considerations to bear in mind regarding the research question. First, the current proposal is conceptually different from the work of Shengelia et al. Second, the procedure is limited in its ability to estimate a causal effect attributable to a healthcare intervention. Third, the definitions of quality are debatable and different in the fields of effective coverage and health services research.

6.3.1. Differences with Shengelia et al.'s approach

Shengelia et al. define effective coverage as an attribute of individuals, while my proposal defines effective coverage as an attribute of healthcare interventions understood as a public policy. This topic was introduced in chapter four.

Shengelia et al. conceive of quality at the individual level, as a latent construct built using the probability of access to different providers, with varying health gains for each one of them. They also assume that providers can have different maximum health gains.

In my proposal, quality is measured using the expected health gain at the individual or population level, and the maximum health gain is extracted from individual predictions, regardless of the healthcare interventions offered by the provider. Consequently, the probability of access to different providers can be prescinded, and it is assumed there is only one maximum health gain.

Shengelia et al. did not discuss the need to use different maximum health gains for each healthcare intervention provider. However, since the maximum health gain is a standard of comparison, it seems reasonable to use one to assess all other providers. This is especially relevant when, according to the framework of Shengelia et al., quality depends on the fulfilment of specific standards of care from healthcare providers.

Here it is important to notice that such a standard – i.e. the maximum health gain – can come from different sources: theoretical considerations from experts, or from the literature, or from following the approach of this thesis, which uses an empirical measurement. Theoretical considerations could originate from experts' opinions that, for instance, fix the minimum health benefit expected from each provider according to feasibility aspects or budget restrictions. The literature can also provide a value for maximum health gain. However, to be consistent with the proposal of this thesis, the outcome should be measured as a continuous variable, such as saved years lived with disability, or its counterpart using health-state utilities. Unfortunately, the scientific literature tends to measure the effectiveness of interventions using disease-specific outcomes.

On the other hand, using an empirical measurement approach has the advantage of providing a feasible standard extracted from the population of interest, including analysis of the specific characteristics of that population and its healthcare providers.

Understanding effective coverage as an attribute of the healthcare intervention is consistent with Tanahashi's original definition of effective coverage, and might be easier to transmit to decision-makers than explaining a non-observed construct at the individual level, as suggested by Shengelia.

Additionally, despite the option of understanding effective coverage as an attribute of healthcare interventions, the proposal of this thesis also allows us to estimate the

components of effective coverage at the individual level. Moreover, chapter four showed that results assuming effective coverage at an individual attribute are similar to those obtained following Shengelia's formulation. However, as mentioned in that chapter, the rationale that supports the proposal of this thesis also supports the idea of effective coverage as an attribute of healthcare.

Finally, in the case of assessing effective coverage for different providers, a stratified analysis can be implemented along with my proposal, though this will reduce the statistical power. Another alternative, as shown in Figure 4.5, is adding the provider variable to the regression model used for predicting disability and then visualising the expected health gain associated with each one, adjusted by covariables.

6.3.2. A causal effect attributable to a healthcare intervention

In deciding how well the thesis answers the research question, a second consideration would be the lack of reliable results for health gain when using data from cross-sectional studies. This issue has been pointed out repeatedly in chapters three, four and five. It is a severe limitation that threatens the research objectives, and therefore it is discussed in more detail in the sections concerning future challenges.

6.3.3. The concept of quality of healthcare interventions

The third consideration to discuss is whether the quality concept used in the effective coverage measurement is coincident or not, with the idea of quality commonly used in health services research.

Unfortunately, there is no single and commonly used concept of quality for healthcare assessment. One of the first approaches to the topic of quality of health services was developed by A. Donabedian, who identified three dimensions for assessment: structure, processes, and outcomes of care. [4]. The proposal of this thesis to measure quality clearly matches the concept of 'outcome of care'. What's more, Donabedian even makes a distinction between different standards of comparison to assess quality, using either an

empirical approach (as in the case of this thesis) or a normative one (such as literature or expert opinion).

Another essential reference is the Institute of Medicine (IoM) from the US, which has exerted a strong influence on how the quality of health systems has been understood within the WHO, and especially in low- and middle-income countries [5]. In 2001, the IoM defined quality as the degree to which a healthcare intervention increases the likelihood of desired health outcomes, in accordance with professional knowledge. Interestingly, for measuring purposes, the IoM agrees with the Donabedian approach [6]. However, that institute recognises six domains (or aims) of quality: safety, effectiveness, the degree to which the provision of care is centred on the patient, timeliness, efficiency, and equity [7]. The measure of quality in this thesis is closest to the perspective of effectiveness. At the same time, safety and timeliness are integrated to the extent that they impact the outcome of disability or loss of health-state utilities. Efficiency and equity are not included in the procedure to calculate the quality of a healthcare intervention in this thesis. However, the convenience of including them into the concept of quality may need to be discussed. It is also worth noting that the indicator of absolute effective coverage may incorporate efficiency when including the budget addressed to the healthcare intervention in its calculations.

More recently, a Lancet Commission focused on the topic of 'high-quality health systems in the era of SDG', defining a high-quality health system as 'one that optimises healthcare in a given context by consistently delivering care that improves or maintains health outcomes, by being valued and trusted by all people, and by responding to changing population needs' [5]. This definition keeps central the idea of maintaining or improving health outcomes, but also includes several other aspects that broaden the concept of quality. However, the overall framework of high-quality health systems remains consistent with Donabedian's framework, since the Lancet Commission recognises three components: structure, process and 'quality impact' (more or less equivalent to care outcomes). Within quality impact, 'better health' is included, which could match with the concept of quality used in this thesis. However, it goes beyond by considering the level of trust in the health system, as well as the financial protection, which clearly escape the scope of what I have proposed here.

Finally, the Organisation for Economic Co-operation and Development (OECD), in its definition of quality of care, includes three dimensions: effectiveness, safety and

responsiveness/patient-centredness [8]. Again, this concept is broader than the one I have used in the framework of effective coverage.

In summary, the definition I have used to measure quality in the context of effective coverage is narrower than the usual conceptualisation of quality in the literature about health systems. However, metrically it seems to be consistent with one of the central aspects of the concept: the extent to which healthcare improves a population's health outcomes.

Notwithstanding this coincidence, it is important to remember that the outcomes of care also depend on aspects originating from patients, such as adherence to treatment. This issue is not usually emphasised when care outcomes and, therefore quality, are discussed.

In conclusion, this thesis answers to a large extent the research question raised in Chapter One. On the one hand, the task of including a component of quality based on the concept of health gain attributable to a healthcare intervention was achieved, although it moved away from the original proposal by Shengelia et al. On the other hand, the concept of quality is narrower than current conceptualisations, and, more importantly, the causal attribution of health gains to a healthcare intervention has not been entirely resolved. This last element is discussed alongside the other challenges for future research.

6.4. Strengths of the proposed procedure to measure effective coverage

The strengths of the procedure have been presented throughout the different chapters of this thesis. However, they are summarised in Table 6.1.

Table 6.1. Main strengths of the proposed procedure to measure effective coverage

	Strengths	Chapter
1	The general procedure is based on calculating attributional fractions on a continuous outcome. It is an **innovative approach** since previous procedures were mainly for dichotomous outcomes [37, 38].	3
2	The procedure is **based on a well-documented framework of health valuation.**	2
3	The general procedure uses **only one source of information** to calculate all parameters that are required (efficiency criterion).	4
4	Using one source of information allows **consistency between parameters**.	3 and 4
5	The procedure is implemented using **data from health surveys, which are routinely conducted** (more often than cohort studies), and frequently have national representativeness.	4
6	The procedure is especially **appropriate to assess non-communicable diseases**, which are commonly neglected in studies of effective coverage.	1
7	As per Shengelia et at., **utilisation is adequately measured** because it is conditioned to the presence of a normative need [2].	1
8	**The estimator is based on the 'amount of need',** rather than the number of people with a normative need. This would be the optimal approach, according to Tanahashi [1].	**1**
9	The procedure **expands the concept of effective coverage** towards other indicators: relative benefit, quality, and relative and absolute effective coverage.	4
10	The procedure allows the estimation of **not only quality but also the effectiveness** of a healthcare intervention.	4 and 5
11	The indicators can be **expressed using fractions (i.e. relative estimates) or absolute numbers.**	3 and 4
12	The absolute effective coverage (a-EC) allows **direct comparison between healthcare interventions** addressing different diseases.	4, 5 and 6
13	The a-EC is **suitable in combination with the cost of the healthcare conditions**, allowing us to calculate a ratio of the cost per unit of avoided disability.	4, 5 and 6
14	The procedure is **highly flexible**, meaning the assumptions about interactions between variables and control by confounders can be explored straightforwardly.	3 and 4
15	The outcome used for effective coverage can be modified for different purposes: **disability, health-state utilities, quality of life**, etc.	3 and 4
16	The procedure can be used **to compare the performance of healthcare interventions across countries**. That is done using a single maximum health gain for all the countries under study.	5

17	The procedure allows the **effective coverage according to different domains of disability** (or health-state utilities) to be explored.	2, 3 and 6
18	The procedure offers an alternative method for **resolving the lack of information about the quality** of healthcare services, which precludes using the effective coverage approach.	4 and 6

6.5. Challenges for future research in effective coverage

6.5.1 A better proxy for the causal effect attributable to a healthcare intervention

The proposal in this thesis was developed by testing the procedures with data from cross-sectional household surveys, which are used routinely by countries to monitor the health status of populations and assess the implementation of social programs, including the coverage of healthcare services. Cross-sectional surveys are less costly than collecting data from other types of study, such as cohort studies, meaning they are more widely used. Additionally, household surveys are especially appropriate for characterising people with chronic non-communicable diseases, which are conditions that have been partially neglected in the study of effective coverage, despite their importance in terms of the burden of diseases and costs for the health system.

However, cross-sectional studies are not well-suited to studying causality, because they cannot guarantee that the exposure occurred before the outcome and they are susceptible to selection and measurement bias. In addition, it is not possible to account for unobserved confounders in their analysis [9, 10]. Furthermore, household surveys are usually underpowered for detecting or studying diseases with low prevalence.

These limitations were not explicitly addressed in the development of the methodology proposed in this thesis, except for the effort of adjusting by observed confounders. They were also addressed in chapter four by exploring the consistency of results throughout different definitions of the exposure to a healthcare intervention.

The most evident consequence of the failure to estimate a causal effect attributable to healthcare interventions was observed in the distribution of the predicted individual health gains, which often included negative values (see chapters four and five). Apart from the effect of selection bias (e.g. people with the worst disabilities died or else did not want to participate, while people who had no access to treatment were more reluctant to participate in a survey), and measurement bias (e.g. people with mild disability are less likely to remember accessing healthcare), the negative health gains can be explained because people who access a healthcare service are usually more disabled and therefore more willing to seek treatment. The procedure of assuming a negative health gain equal to zero, used in chapters

four and five, is questionable. However, it was implemented to allow discussion about the general aspects of the procedure and interpretation of the different indicators.

Unfortunately, given the nature of the study design, it will never be possible to calculate a robust estimation of the causal effect of a healthcare intervention using cross-sectional data. However, several methodological approaches allow us to make an estimate closer to the real size of the effect of the healthcare intervention and are used routinely in the analysis of surveys.

Propensity scores are an attractive alternative that deserve to be investigated in the future. A propensity for utilising a healthcare intervention can be calculated with a logistic regression model, including a similar set of covariates to those already incorporated into the model to estimate disability or health-state utilities. Then, the propensity score can be used to estimate the effect of the healthcare intervention through different methods, specifically stratification, matching, weighting, and adjustment as a regressor [11].

Classical stratification implies estimating the effect size of a treatment using people exposed and not exposed from the same propensity quantile [12]. Matching means estimating the effect size by comparing the number of exposed with a fixed number of unexposed, selected according to proximity rules in their propensity [13]. This procedure usually discards several unexposed cases from the analysis. The weighting family (i.e., inverse propensity, fine stratification weights, matching weights, overlap weights) uses propensity scores to weight the outcome from each individual, ensuring balance in their propensity according to its distribution among those exposed or among the whole population [11]. Finally, the propensity score can be used as an additional variable in the regression model used to calculate health gains.

Each alternative has pros and cons [14], and they can be implemented with relative ease to calculate the r-EC and the a-EC (relative and absolute effective coverage, respectively). However, in the case of the indicator of effective coverage itself, it is necessary to generate a distribution of health gain based on individual predictions. Unfortunately, the procedures described above are usually implemented to obtain only average estimates of treatment effects. Among the propensity score procedures, adding the propensity as a regressor seems the simplest way to resolve this issue. However, some studies have shown that including the

propensity as a regressor appears not to improve the estimation of the treatment effect substantially [14].

Propensity score matching has been used previously to estimate effective coverage, although it makes use of health outcomes different from disability or heath-state utilities [15, 16].

Another alternative commonly used to resolve residual confusion as well as endogeneity between the dependent and independent variables is the use of instrumental variables. [17]. On the relationship between a treatment and a health outcome, an instrumental variable correlates with the health outcome only through the treatment [17, 18]. However, finding a variable that complies with the requirements for a good instrumental variable is often challenging, especially in cross-sectional studies.

In the context of effective coverage, some authors have followed the instrumental variable approach [19], where attributes of the population-level are usually chosen [20]. For instance, the coverage of a vaccine at a district level is correlated with the access to that vaccine at the individual level, but not necessarily with all causes of mortality [19]. Or, another example, the regional rate of cardiac catheterisation is correlated with the individual access to such an intervention in people with acute myocardial infarction, but not with mortality in the follow-up [21]. Fortunately, the assumptions behind the selection of an instrumental variable are susceptible to confirmation.

In the context of the surveys analysed in this thesis, the coverage for hypertension, depression and osteoarthritis at the level of sub-national regions might be explored as a potential instrumental variable. When information about the sub-national regions is lacking, the sample unit could be used.

Using an instrumental variable approach at a population level might carry the challenge of adapting the procedure to estimate health gains using multilevel approaches, such as random effects. This is especially complex in contexts where the data is weighted according to a specific sample design. Regardless of this complexity, the procedure can be suitable for producing individual predictions to estimate effective coverage. Moreover, using a multilevel approach allows adjustment for unobserved factors from the population level [15, 22].

A third alternative would be using regression discontinuity designs. This is a procedure based on regression models that calculates the treatment effect among people close to a specific cut-off, according to a scale that defines who receives the healthcare intervention [23]. For example, in the case of depressive disorder, as explored in chapters four and five, the need for treatment is normatively defined by a score above four on a scale of symptoms. The regression discontinuity design proposes calculating the treatment effect by comparing people whose scores are below or above the threshold. This approach assumes the interchangeability of the characteristics of people with and without treatment when they are close to the cut-off point. It is implementable for hypertension, diabetes, and other diseases, where a continuous scale defines the need for treatment. The approach also justifies the comparison arguing a random error in measuring the scale, which can work as a form of 'local randomisation'.

However, this procedure has the limitation that it applies only when utilisation is a function of rules based on a continuous scale. Furthermore, it is challenging to implement for predicting individual health gains, as required to estimate effective coverage. Moreover, the procedure only calculates the treatment effect for those cases with diseases in stages close to the cut-off point. However, in chapter four, when looking at depressive disorders, I showed that the health gains associated with the healthcare intervention differed notably in different areas depending on the scale of symptoms.

In summary, several approaches must still be investigated, and some of them might improve the estimation of the health gain attributable to healthcare interventions.

6.5.2 Including fatal consequences in effective coverage

According to this thesis, effective coverage is measured by considering the quality of healthcare service (i.e. heath gain/ maximum heath gain), while r-EC and a-EC are measured by considering the health benefit associated with such healthcare services (i.e. health gain/ disability attributable to a disease). Therefore, to include the fatal consequences of diseases in the estimation of effective coverage, it would be necessary to add them to each one of these parameters: heath gain, maximum health gain, and the disability (or health-state utilities) attributable to the disease. For this, it would be necessary to use the information on

death registers coming from the same target population of the survey used to calculate effective coverage.

In the case of the disability attributable to a disease, it seems less complicated to add the fatal consequences, as long as the scale used to measure disability is anchored in the state of death (see chapter two). This is the case when the outcome of measuring health gains is health-state utilities. When disability is used, it must be assumed that the extreme level of disability is equivalent to death. In such cases, the number of years lost due to death by the disease occurred before a standard of life expectancy could be added directly to the total number of person-years of disability (or health-state utilities) [24].

This approach is no different from the current methods used to calculate disability-adjusted life years (DALY), where the number of years lived with a disability (YLD) are added to the years of life lost by premature death (YLL).

To add fatal consequences to health gains is much more challenging. Health gains are calculated by contrasting the current scenario (i.e. the years of life lost attributable to a disease) against a counterfactual scenario where no one in need uses the healthcare service; or, its equivalent, where the effectiveness of the treatment is zero. To estimate how many deaths are avoided in people who use the healthcare service, it would be necessary for the certificate of death to identify whether or not the person in question was able to utilise that service.

Another alternative would be to assume that the relative benefit calculated for disability is equivalent to the impact of the healthcare service on avoiding death. The assumption of the interchangeability of the effect size of an exposure on the incidence of disability and mortality has been used routinely in studies of comparative risk assessment [25]. In other words, in the assessment of risk factors, the same relative risk is used to calculate the YLD and YLL.

Under that assumption, the 'years of death avoided' or 'years of life gained' by the healthcare intervention would be equivalent to:

Years of life gained = Odds (relative benefit) * YLL by the disease [Eq. 6.1]

Using ‘years of life gained”, the r-EC, including fatal consequences (using the population-average level notation), would be expressed as:

$$\text{r-EC} = \frac{Ut*HG_{G2} + \text{years of life gained}}{(1-Ut)*DA_{G1} + UtDA_{G2}' + YLL} \qquad \text{[Eq.6.2]}$$

and the a-EC would be expressed as::

$$\text{a-EC} = \frac{Ut*NN*HG_{G2} + \text{years of life gained}}{NN\,[D_{G1}\,(1-Ut) + D_{G2}''Ut] + [D_{G0}(1-NN)] + \text{total YLL}} \qquad \text{[Eq.6.3]}$$

where Ut is utilisation of a health service; NN is normative need; HG is health gains; DA is attributable disability; G0 is people without a NN; G1 is people with a NN but without utilisation; and G2 is people with a NN and utilisation (see chapter four, Table 4.1. for further details).

On the other hand, it is necessary to calculate the maximum avoidable number of deaths to estimate effective coverage, which would be equivalent to the maximum health gain. Following the approach presented above, a possible approximation would come from calculating the fraction of avoidable disability over the total disability attributable to disease, and then applying it to the YLL.

That fraction of avoidable disability (i.e. FAD) can be calculated: FAD = (DA – HGmax) / DA, where HGmax is the maximum health gain.

Consequently, the effective coverage, including fatal consequences, would be expressed (using a population-average level notation) as:

$$\text{Effective coverage} = \frac{Ut*HG_{G2} + \text{years of life gained}}{HGmax_{G2} + YLL*FAD} \qquad \text{[Eq.6.4]}$$

See chapter four, Table 4.1. for further details about the original formulation of effective coverage.

These proposals need to be evaluated carefully in future research. The assumptions behind these formulations need to be studied in more detail. Here, I have shown only an initial approach that could be used as a point of departure for enriching the metric of effective coverage.

6.5.3 Including future consequences of diseases

Another relevant aspect in the metric of effective coverage is the inclusion of the impact of the healthcare services on the future consequences of a given disease.

Under the approach of this thesis, the effective coverage measures the amount of the current disability attributable to a disease that is being avoided by the utilisation of healthcare. This is also valid in cases where one wishes to incorporate the fatal consequences, since we are calculating the YLL that occurred at the moment of measuring effective coverage. However, the disability or deaths that will be avoided in the future by receiving treatment now is not considered in the analysis.

This is partially true. When we estimate the effect of the healthcare services, for instance, on hypertension, we include in the analysis people at different stages and with different trajectories in the disability associated with this disease. It is likely that, among people with hypertension but without treatment, the disability associated with the adverse consequences of hypertension will be higher than for people with treatment. However, some of the people without treatment could have died prematurely, hiding the excess of disability. Considering future consequences in this way assumes a stable population in terms of demographics and the use of health services.

On the other hand, following a cross-sectional perspective in assessing disability is the traditional approach for burden of disease studies. In those studies, the burden of hypertension is measured separately from its consequences (coronary diseases, stroke, etc.).

Therefore, the approach suggested in this thesis could be especially valid in the case of assessing curative and rehabilitative healthcare interventions. That said, it could be weaker for assessing preventive healthcare services. However, this is not an inherent limitation of the procedure, but only a consequence of relying on cross-sectional data.

This might be a significant limitation, since some relevant healthcare interventions are preventive ones: vaccination programs, antenatal care, basic sanitisation, cancer screening, among others. This approach may not be the most appropriate for comparing health gains and disability between people who are vaccinated and not vaccinated, or people with and without breast cancer screening, since the incidence and duration of the consequences cannot be assessed reliably through cross-sectional studies.

In summary, the approach suggested in this thesis partially includes the future consequences of disability using a cross-sectional approximation, although with some underlying assumptions behind it. However, the cross-sectional approach precludes assessing other common preventive healthcare interventions.

6.5.4. Including an equity perspective

Equity in the distribution of health outcomes has usually been considered a central goal of health systems, attempting simultaneously to improve the average level of health, the financial protection, and the responsiveness of healthcare services, among other aims [26, 27]. If effective coverage pretends to be a valuable instrument to assess the performance of health systems, it should be able to include such a perspective.

However, equity aspects are often lacking in the metrics of effective coverage. For example, neither Tanahashi nor Shengelia et al. include a formal proposal in this area. Moreover, in the systematic scope review presented in chapter one, very few of the publications had equity considerations when assessing effective coverage.

How to use or adapt the effective coverage measurement to allow increasing insights into the health inequalities behind healthcare interventions is one of the pending challenges that should be addressed in the future.

Several approaches can be followed. One obvious alternative is carrying out a stratified analysis of effective coverage by groups of interest (educational level, quantile of income, gender, ethnicity, type of provider [public or private], etc.). However, the limit of this approach is the loss of statistical power to perform inferences, producing estimates with high uncertainty, or even precluding the implementation of parametric procedures.

Another alternative, that was alluded to earlier in this thesis and is perhaps more interesting, would be exploring the health gains attributable to the healthcare intervention stratified by the group of interest. In chapter four (see Figure 4.4), examples in this regard are presented. It is worth remembering that differences in health gains, observed by educational level or income, were adjusted by the presence of other covariables (i.e., marital status, comorbidities, etc.). Therefore, inequalities can be attributable to differences in the quality of the healthcare intervention.

However, there is another approach that could be even more interesting to explore in future research, although a little more complex.

Chapter three presents a straightforward approach for measuring the disability (or health-state utilities) attributable to a disease. Using that same approach, it is possible to estimate the disability associated with an individual's social position, adjusting or not adjusting by the presence of diseases. Then, using a procedure similar to the one implemented to calculate effective coverage, we can estimate how much of the disability associated with that social position is due to the coverage and quality (or lack of quality) of different healthcare interventions. Again, this can be calculated among people with a particular disease or in the population as a whole.

In other words, it could be possible to calculate not only how much disability can be avoided via increasing the coverage and the quality of healthcare services, but also calculate how much of the disability that comes from social inequalities could be avoided (or increased). Moreover, because the indicators can be expressed in units of disability valid for any disease, it would be possible to compare healthcare interventions – increasing the average health of the population and diminishing health inequalities, in terms of efficiency – especially when the cost of each unit of disability is included in the analysis.

The procedures developed in this thesis might be a powerful tool to evaluate, using a single framework of assessment, the impact of different scenarios of healthcare interventions, both in health improvements and the equity of the distribution of such health improvements.

The relationship between effective coverage and health inequalities associated with healthcare services seems an interesting field for developing further insights into the overall performance of health systems.

6.5.5. Including an economic dimension

In chapters four and five, I mentioned that the indicator of a-EC is suitable for combining with the costs of healthcare interventions through an expression of cost per unit of health consequences.

The indicator of a-EC is especially suitable because it allows the performance of different healthcare interventions to be measured using the same metric unit: the avoided fraction of the total population disability (or health-state utilities). The avoided fraction can also be expressed in absolute terms, through the total person-years of avoided disability.

The possibility of calculating the cost of each unit of absolute effective coverage offers an interesting path for future research. In this regard, several elements might be worth exploring.

First, the convenience of the outcome used to measure absolute effective coverage. In chapter two, I argued that health-state utilities might be more suitable than disability for informing decision making. That is because utilities include the social preferences for each health state, which is consistent with theories for allocating resources [28, 29].

Second, it is important to decide how to calculate costs. Contrary to the traditional approaches used in health technology assessments, where the costs and benefits are interrelated and defined a priori by a model, in the approach of this thesis, the benefits are obtained from observations, while the costs are calculated separately. Thus, one approximation for calculating costs could be estimating the price of an average basket of health benefits associated with a healthcare intervention. However, since there are different

perspectives for estimating costs (e.g. a perspective from the provider, payer, patient, whole society), each one might deserve a different approach [30]. Moreover, as mentioned above, since the health gains are calculated by integrating the disability of people with different utilisation trajectories, the cost should also include those trajectories of utilisation for each component of the basket of health benefits. This exercise is done regularly by public and private insurance companies [31].

On the other hand, it is tempting to interpret the costs by health gains equivalent to DALYs and the quality-adjusted life years (QALYs) in cost-effectiveness studies. In chapter three, I argued that the disability attributable to a disease was equivalent to calculating YLDs. Therefore, the health gains from effective coverage might be equivalent, at least conceptually, to the avoided YLDs attributable to the utilisation of a healthcare intervention. Furthermore, in that chapter, I mentioned that DALYs and QALYs were calculated by multiplying disability weights (or health-state utilities) by the time lived in each health state [32, 33]. Besides, the main difference between the usual way of calculating DALYs and QALYs, and the way proposed in this thesis, is the calculation for the time lived in each state. The cross-sectional approach used here is consistent with the standard procedure for calculating DALYs. However, to the best of my knowledge, it has not been documented in the case of QALYs, except for those articles I published in the context of this thesis [34, 35].

Third, if the theory for decision making that justifies cost-effectiveness studies was transferred to the context of effective coverage estimation, it would generate rich opportunities for expanding the tools available to support the decision-making process. This would be especially significant if we consider that, under the approach of this thesis, the analyses are carried out using data from nationally representative health surveys.

Additionally, these considerations could be even more interesting if the fatal consequences of diseases could be added to the metric of effective coverage, and a standard equity perspective were implemented. Through this approach, an important step in studying the classical trade-off between maximising health and maximizing an equitable distribution of health could be made [36].

6.5.6 A brief consideration of the 'health gain threshold'.

Another relevant challenge for future research is a deeper assessment of the role played by the threshold used for determining a standard of comparison to calculate the quality of healthcare services. This issue was not exhaustively evaluated in this thesis, but there are two main approximations for performing the appraisal. One is a practical way, where the behaviour of quality and effective coverage is explored against different selected thresholds. This can be accomplished using real data, or, even better, by simulating populations under different assumptions.

The second approximation is a theoretical one. In chapter four, I discussed how using a threshold implies the assumption that all differences in health gains under the selected value are due to differences in the quality of the healthcare intervention (i.e. adherence and following treatment standards). This assumes that treatment effectiveness does not vary (i.e. does not interact) with individual attributes such as age or sex – something that needs to be proven.

Choosing an appropriate threshold will depend on both empirical and theoretical considerations. Both require more investigation.

6.4.7 Analysis sub-domain of disability (or health state utilities)

Finally, in the second chapter, I showed that first-order latent variables, corresponding to different sub-domains of disability, might be calculated using a structural equation model on the questionnaire that was analysed. The same approach was used for the Health State Description questionnaire, a disability questionnaire from the SAGE study introduced in chapter five. In addition, calculating the disability attributable to various disability sub-domains was shown in chapter three. Moreover, the EQ5D is also susceptible to being decomposed by sub-domains of health-state utilities [34, 35].

Consequently, the outcome of effective coverage can also be decomposed by disability sub-domains (or health-state utilities). This feature is very interesting, because the performance of healthcare services can be assessed more precisely according to different therapeutic aims.

For example, in the case of osteoarthritis, effective coverage could be expressed separately for the sub-domain of pain and the sub-domain of mobility. In the case of depressive disorder, the sub-domains of mood and energy might be explored apart from other sub-domains. Or, if there is a justified underlying hypothesis, the avoided disability attributable to (physical) pain can be calculated in people that receive treatment for depression.

The multidimensionality of the outcome is potentially a valuable topic related to effective coverage that emerges from this thesis and that deserves more investigation.

In conclusion, there are still several knowledge gaps and challenges regarding the metric of effective coverage and its related indicators. The main focus of this thesis was to explore multiple alternative approaches to implementing an overall procedure that would include health gains consistently with the health-state valuation framework. The approach presented in this thesis will need to be improved as we address the many challenges outlined above, but at least the topics for future research are well identified.

6.6. Is effective coverage a good indicator for tracking progress toward Universal Health Coverage (UHC)?

Up to this point, I have discussed how the thesis's research question has been answered, along with the strengths of the proposal and any challenges for future research. In this section, I will briefly discuss a topic beyond the strict remit of the research aims: that is, whether this new proposal to calculate effective coverage is suitable for monitoring UHC. This is relevant because several groups have claimed that effective coverage, under Shengelia et al.'s framework, is the most appropriate indicator for such a goal [39-41]. However, none of them has been able to implement it adequately.

In chapter one, it was mentioned that UHC has three dimensions: population (who is covered), services (which services are covered), and direct cost (proportion of the cost covered). Additionally, some frequent criticisms made against the UHC movement were pointed out: (1) coverage is understood as mere accessibility, without including the opportunity of access or the quality of the healthcare services; (2) UHC is mainly centred on how to finance healthcare services, rather than the efficiency, quality, and effectiveness of those services; and (3) the approach to UHC neglects the role of extra-sectorial actors. Also, I can add that the lack of information about the quality of healthcare interventions has precluded implementing the framework proposed by Shengelia et al. [41-43].

6.6.1 Dimensions of UHC and effective coverage

Regarding the dimensions included in the concept of UHC, the procedure proposed in this thesis can clearly address the first of them: the population covered. The proposed approach quantifies the fraction of people with a normative need that receive the healthcare intervention. Additionally, it quantifies the fraction of their disability attributable to the disease which is avoided by the healthcare service. Conceptually, using disability or health-state utilities might be a better outcome for decision making than just the number of people with a normative need, as is proposed by the UHC framework.

The second dimension for which services are covered is considered by the procedure raised in this thesis. One way to integrate it is by adding effective coverage from different healthcare interventions, as is currently proposed by the institutions monitoring the UHC.

How to add the effective coverage from different healthcare interventions into a single expression that can summarise the performance of health services was beyond the scope of previous chapters. However, it is not a difficult task. Shengelia et al. proposed weighting each effective coverage by the sum of the expected maximum health gains from all healthcare interventions. That is a reasonable alternative implementable under the framework of this thesis (see Table 6.2 for an example).

Table 6.2. Example for summarising effective coverage from different healthcare interventions.

According to chapter five, in China, the effective coverage for depressive disorder, hypertension and osteoarthritis was 10.3%, 5.1% and 13.0%, respectively (see Tables 5.6, 5.9, 5.12). In addition, the maximum health gain potentially achievable for those conditions assuming full coverage and full quality was 135,571 person-years, 2,162,493 person-years, and 732,283 person-years, respectively (see HG'_{G1} + HG'_{G2} in Tables 5.5, 5.8, 5.11). A total effective coverage can be obtained by summing them, weighting for the fraction that each maximum health gain represents on the total maximum health gains: (10.3% * 0.044) + (5.1% * 0.714) + (13.0% * 0.242) = 7.24%.

In the case of r-EC, a similar procedure can be implemented. However, instead of weighting each indicator by the maximum health gain, the disability attributable (DA_1 + DA_2) should be used. The case of the a-EC indicator is even simpler. They do not require any weight and can be added up, as long as it is assumed that the effect of each healthcare intervention will not substantially change the total level of disability in the whole population. Otherwise, it can be derived using the following equation:

$$\text{a-EC} = \frac{\sum_j^n Odds\ (aEC_j)}{1 + \sum_j^n Odds\ (aEC_j)} \qquad \text{[Eq. 6.5]}$$

where *aEC* is the absolute effective coverage, *j* is the healthcare intervention, and *n* is the number of healthcare interventions whose absolute effective coverage must be added.

Recently, the Institute for Health Metrics and Evaluation (IHME) in the US attempted to create an effective coverage indicator closer to the Shengelia et al. framework, proposing to weight each sub-indicator of UHC (see Table 1.3) by disability weights extracted from the burden of disease study for each disease [44]. However, that approach does not seem entirely coherent with the framework of Shengelia et al., because it equates the concept of disability weights (or attributable disability) with the idea of quality (or maximum health gain). On the other hand, it is worth remembering that the strategy proposed by the WHO and the World Bank is not to weight any sub-indicators, assuming they all have equivalent weights [42] (see chapter one).

In conclusion, since the procedure proposed in this thesis overcomes the limitation of the lack of a quality indicator, and their resultant indicators can be added up, effective coverage, according to Shengelia et al., might be empirically repositioned as an alternative for measuring UHC.

However, there is a more subtle way to include the extent of health services covered by the health system, the second dimension of the UHC. That is by considering the total disability from the whole population. In that context, it is valuable to remember that, under Shengelia et al.'s framework, it is impossible to know the total need of the population, just the total avoidable disability that depends on the healthcare interventions analysed. In addition, for practical reasons, the proposals from the IHME and the WHO, jointly with the World Bank, selected just a few healthcare interventions to be integrated into the final indicator of UHC. But how much of the total disability in each country is resolved, or how much health is produced given the coverage of those selected healthcare interventions, is not discussed.

Using the total disability of the entire population as a target for health services might be more convenient than using the population coverage and the number of health services, as the framework of UHC proposes. The total population disability (or the health-state utilities lost) integrates into a single expression the health consequences from the number of people covered and the number of healthcare services offered. This approach is more straightforward and allows the trade-off between increasing the coverage of a healthcare intervention or increasing the number of interventions to be shown in a better way.

The a-EC offers an interesting answer in this regard, mainly because it allows the impact of each health service to be compared directly through a single reference parameter: the total disability of the entire population. This property makes the a-EC indicator much more comprehensive than the indicators currently used for monitoring UHC.

The third dimension for measuring UHC is more challenging to address. People's financial protection has also been considered a central goal of health services [26, 27]. However, its metric is based on different criteria than disability or health-state utilities. One option is measuring the utilities associated with each year of financial protection, to combine with the outcome of effective coverage. However, the convenience of producing one measure that integrates all aspects of UHC must be discussed further.

6.6.2 Criticisms of UHC and effective coverage

Regarding the criticisms addressed to the UHC movement, the metrics proposed in this thesis can at least provide answers to some of them.

Firstly, the way this thesis proposes measuring effective coverage allows us to include some elements of access beyond mere accessibility. Thus, the quality of healthcare services is integrated, as well as the opportunities for access, at least to the extent that this affects the health gains and the disability of people concerned.

Secondly, r-EC and a-EC indicators combine the effectiveness of the healthcare intervention with its coverage. On the other hand, the effective coverage indicator combines quality with coverage. Furthermore, because the a-EC can potentially be used to estimate the cost per unit of avoided disability (or health-state utility produced), it might be suitable to assess the distributive efficiency across different healthcare interventions. In other words, the metric proposed in this thesis allows the criteria of quality, effectiveness, and efficiency to be included in the analysis of UHC.

Finally, the indicator of a-EC includes not just the avoidable disability by a few healthcare interventions, but the total disability of the population. In addition to the parameters for effectiveness and coverage, that indicator also includes the prevalence of health conditions, which allows, in a certain way, the burden from underlying factors that generate the disease

to be considered. Therefore, the impact of healthcare interventions from all other causes that produce disability – potentially including the role of determinants of health – can be weighted using this approach.

In summary, the proposal of this thesis offers an alternative for generating the parameters of quality of the healthcare services, enabling Shengelia et al.'s framework for measuring effective coverage to be implemented. However, in the work of this thesis, a-EC arises as a potentially more suitable indicator for monitoring the progress of UHC.

Absolute effective coverage (a-EC) simplifies the framework of the UHC, highlighting the fact that the final goal of health systems is to improve the health of people, and that access to healthcare services is only one of these paths. The a-EC can be increased by improving the effectiveness, quality, coverage, and diversity of healthcare interventions. Furthermore, the a-EC can be improved by decreasing the prevalence of those diseases with the highest disability. Moreover, it does not require special weights to create summary measures, and it can be combined with the cost of healthcare interventions informing the decision-making process.

6.7. Conclusion

This thesis achieves the aims postulated in its formulation: that is, to develop a new approach to include the quality component in the estimation of effective coverage using health gains. However, this proposal still has several limitations, which are interesting topics for future research.

The results of this thesis correspond to an expanded version of effective coverage developed by Tanahashi and Shengelia et al. Several new indicators have been proposed, which are useful for assessing the performance of the health services. Among them, the a-EC seems a suitable alternative for monitoring UHC progress.

References of chapter six

1. Tanahashi T: **Health service coverage and its evaluation**. *Bulletin of the World Health Organization* 1978, **56**(2):295-303.
2. Shengelia B, Tandon A, B. Adams O, Murray CJL: **Access, utilization, quality, and effective coverage: An integrated conceptual framework and measurement strategy**. *Social Science & Medicine* 2005, **61**:97-109.
3. United Nations: **Transforming our World: The 2030 Agenda for Sustenaible Development**. In: *A/RES/70/1.* Edited by Nations U: United Nations; 2015.
4. Donabedian A: **Evaluating the quality of medical care. 1966**. *Milbank Q* 2005, **83**(4):691-729.
5. Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, Adeyi O, Barker P, Daelmans B, Doubova SV *et al*: **High-quality health systems in the Sustainable Development Goals era: time for a revolution**. *The Lancet Global health* 2018, **6**(11):e1196-e1252.
6. Institute of Medicine Committee on Quality of Health Care in A. In: *Crossing the Quality Chasm: A New Health System for the 21st Century.* edn. Washington (DC): National Academies Press (US)

Copyright 2001 by the National Academy of Sciences. All rights reserved.; 2001.

7. Wolfe A: **Institute of Medicine Report: Crossing the Quality Chasm: A New Health Care System for the 21st Century**. *Policy, Politics, & Nursing Practice* 2001, **2**(3):233-235.
8. OECD: **Caring for Quality in Health. Lessons Learnt from 15 Reviews of Health Care Quality**. In: *Reviews of Health Care Quality.* Paris; 2017.
9. Sedgwick P: **Ecological studies: advantages and disadvantages**. *Bmj* 2014, **348**:g2979.
10. Sedgwick P: **Bias in observational study designs: cross sectional studies**. *Bmj* 2015, **350**:h1286.
11. Desai RJ, Franklin JM: **Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners**. *Bmj* 2019, **367**:l5657.
12. Rosenbaum PR, Rubin DB: **The Central Role of the Propensity Score in Observational Studies for Causal Effects**. *Biometrika* 1983, **70**(1):41-55.
13. Austin PC: **An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies**. *Multivariate Behav Res* 2011, **46**(3):399-424.
14. Elze MC, Gregson J, Baber U, Williamson E, Sartori S, Mehran R, Nichols M, Stone GW, Pocock SJ: **Comparison of Propensity Score Methods and Covariate Adjustment: Evaluation in 4 Cardiovascular Studies**. *J Am Coll Cardiol* 2017, **69**(3):345-357.
15. Farzadfar F, Murray CJ, Gakidou E, Bossert T, Namdaritabar H, Alikhani S, Moradi G, Delavari A, Jamshidi H, Ezzati M: **Effectiveness of diabetes and hypertension management by rural primary health-care workers (Behvarz workers) in Iran: a nationally representative observational study**. *Lancet* 2012, **379**(9810):47-54.

16. Eisele TP, Larsen DA, Anglewicz PA, Keating J, Yukich J, Bennett A, Hutchinson P, Steketee RW: **Malaria prevention in pregnancy, birthweight, and neonatal mortality: a meta-analysis of 32 national cross-sectional datasets in Africa**. *Lancet Infect Dis* 2012, **12**(12):942-949.
17. Pokropek A: **Introduction to instrumental variables and their application to large-scale assessment data**. *Large-scale Assessments in Education* 2016, **4**(1):4.
18. Greenland S: **An introduction to instrumental variables for epidemiologists**. *Int J Epidemiol* 2000, **29**(4):722-729.
19. Wong K, Campitelli MA, Stukel TA, Kwong JC: **Estimating influenza vaccine effectiveness in community-dwelling elderly patients using the instrumental variable analysis method**. *Arch Intern Med* 2012, **172**(6):484-491.
20. Maciejewski ML, Brookhart MA: **Using Instrumental Variables to Address Bias From Unobserved Confounders**. *Jama* 2019, **321**(21):2124-2125.
21. Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ: **Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods**. *Jama* 2007, **297**(3):278-285.
22. Diez-Roux AV: **Multilevel analysis in public health research**. *Annu Rev Public Health* 2000, **21**:171-192.
23. Moscoe E, Bor J, Bärnighausen T: **Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice**. *J Clin Epidemiol* 2015, **68**(2):122-133.
24. Murray CJ: **Quantifying the burden of disease: the technical basis for disability-adjusted life years**. *Bull World Health Organ* 1994, **72**(3):429-445.
25. J.L. MC, Majid E, D. LA, Anthony R, Stephen VH: **Comparative quantification of health risks: conceptual framework and methodological issues. Global and Regional Burden of Disease Attributable to Selected Major Risk Factors**. In: *Comparative Quantification of Health Risks. Volume 1*, edn. Edited by Majid E, D LA, J.L. MC. Geneva: World Health Organization; 2004: 1-38.
26. Murray CJL, Frenk J: **A framework for assessing the performance of health systems**. *Bulletin of the World Health Organization* 2000, **78**(6):717-731.
27. Roberts MJ, Hsiao W, Berman P, Reich MR: **Goals for Evaluating Health Systems**. In: *Getting Health Reform Right A Guide to improving Performance and Equity.* edn.: Oxford University Press; 2008.
28. Torrance GW: **Social preferences for health states: An empirical evaluation of three measurement techniques**. *Socio-Economic Planning Sciences* 1976, **10**(3):129-136.
29. Torrance GW: **Toward a utility theory foundation for health status index models**. *Health Serv Res* 1976, **11**(4):349-369.
30. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, GW T: **Methods for the Economic Evaluation of Health Care Programmes**, 4th edn. Oxford, UK: Oxford University Press; 2015.
31. Chile MdSd: **Informe Final del Estudio de Verificación del Costo Esperado Individual Promedio por Beneficiario de Conjunto Priorizado de Problemas de Salud con Garantías Explicitas 2018**. In.; 2019.

32. Murray CJ, Ezzati M, Flaxman AD, Lim S, Lozano R, Michaud C, Naghavi M, Salomon JA, Shibuya K, Vos T *et al*: **GBD 2010: design, definitions, and metrics**. *Lancet* 2012, **380**(9859):2063-2066.
33. Weinstein MC, Torrance G, McGuire A: **QALYs: the basics**. *Value Health* 2009, **12 Suppl 1**:S5-9.
34. Zitko P, Bilbeny N, Balmaceda C, Abbot T, Carcamo C, Espinoza M: **Prevalence, Burden of Disease, and Lost in Health State Utilities attributable to Chronic Musculoskeletal Disorders and Pain in Chile**. *BMC Public Health* 2020, **(under second review)**.
35. Zitko P, Slachevsky A, Jumenez D: **The prevalence, associated factors, and the burden on health state utilities for dementia in Chile**. *Alzheimer's and Demen* 2020, **16**(Suppl. 10):e042466.
36. Cookson R, Mirelman AJ, Griffin S, Asaria M, Dawkins B, Norheim OF, Verguet S, J. Culyer A: **Using Cost-Effectiveness Analysis to Address Health Equity Concerns**. *Value in Health* 2017, **20**(2):206-212.
37. Tanuseputro P, Perez R, Rosella L, Wilson K, Bennett C, Tuna M, Hennessy D, Manson H, Manuel D: **Improving the estimation of the burden of risk factors: an illustrative comparison of methods to measure smoking-attributable mortality**. *Population health metrics* 2015, **13**:5.
38. Palazzo C, Yokota RTC, Ferguson J, Tafforeau J, Ravaud JF, Van Oyen H, Nusselder WJ: **Methods to assess the contribution of diseases to disability using cross-sectional studies: comparison of different versions of the attributable fraction and the attribution method**. *Int J Epidemiol* 2018.
39. Boerma T, AbouZah C, Evans D, Evans T: **Monitoring Intervention Coverage in the Context of Universal Health Coverage**. *PLOS Medicine* 2014, **11**(9).
40. World Health Organization: **Tracking universal health coverage: first global monitoring report**. In.; 2015.
41. Ng M, Fullman N, Dieleman JL, Flaxman AD, Murray CJL, Lim SS: **Effective Coverage: A Metric for Monitoring Universal Health Coverage**. *PLOS Medicine* 2015, **11**(9).
42. World Health Organization and the International Bank for Reconstruction and Development / The World Bank: **Tracking universal health coverage: 2017 global monitoring report**. Switzerland.; 2017.
43. European Observatory on Health Systems and Policies: **Health system performance comparison. An agenda for policy, information and research**: Open University Press; 2013.
44. **Measuring universal health coverage based on an index of effective coverage of health services in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019**. *Lancet* 2020, **396**(10258):1250-1284.

Supplementary material chapter 1

Table S1.1 Electronic search strategy for the scope review on effective coverage

17th April 2020

	Criteria	Number of records	Database
1	Effective.ab,kw,ti.	**3,483,196**	
2	Coverage.ab,kw,ti.	**299,355**	
3	(Effective adj Coverage).ab,kw,ti.	**860**	
4	remove duplicates from 3	**471**	
		97	Embase <1974 to 2020 Week 16>
		304	Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily <1946 to April 17, 2020>
		70	Global Health <1973 to 2020 Week 15>

Table S1.2 Structure of the spreadsheet used to extract information from selected articles in the scope review.

Information	Description
Id	A correlative number
Authors	First author
Year of publishing	Year
Countries under study	Country or list of countries
Target intervention	-
Utilization conditioned to a normative need?	Yes/ No/ not clear
Quality is understood as the gap between observed health gain and the maximum heath gain potential achievable?	Yes/ No/ not clear
Study design	Cross-sectional, cohort, mathematical simulation, etc.
Comments about how the concept of 'Utilization' is addressed	-
Comments about how the concept of 'Quality' is addressed	-
Comment about the Design	-
Additional comments	-
Availability was measured	Yes/ No
Accessibility was measured	Yes/ No
Affordability was measured	Yes/ No
Acceptability was measured	Yes/ No
Contact was measured	Yes/ No
Effectiveness/quality was measured	Yes/ No
Discarded	Yes/ No
Reason why was discarded	-

Table S1.3. Number of studies by sub-type of **maternal, neonatal and general child-care** interventions, that met criteria of utilization according B. Shengelia et al, type of measurement of health gain (quality) and number of countries involved in the assessment of effective coverage.

Intervention	Number of countries were HIC	Number of countries	Number of articles	Number of interventions assessed	**Utilization** is conditioned to a true need	Q1	Q2	Q3	Q4	Q5	Q6
Family planning	0	8	**3**	**3**	3	3	0	0	0	0	0
Antenatal care (general)	1	93	**18**	**18**	16	3	2	12	0	0	0
Antenatal care (18 specific activities) †	0	21	**11**	**32**	18	4	0	27	1	0	0
Obstetric care	0	2	**2**	**2**	2	2	0	0	0	0	0
Skilled birth attendance	1	21	**15**	**17**	14	9	2	4	0	0	0
Postnatal care	1	16	**10**	**10**	9	3	3	3	0	0	0
Preventive health control (child)	1	6	**1**	**1**	1	-	-	-	-	-	-
Child health services	0	1	**2**	**2**	2	2	0	0	0	0	0
Maternal and child health service	1	2	**2**	**2**	2	0	0	1	0	0	0
	1	96	**41**	**87**	77.0%	29.9%	8.0%	54.0%	1.1%	0.0%	0.0%

Q1: Quality type 1 (health care provider)/ Q2: Quality type 2 (incidence of a disease or condition)/ Q3: Quality type 3 (content of heath care)/ Q4: Quality type 4 (biomarker and function)/ Q5: Quality type 5 (activity/ participation)/ Q6: Quality type 6 (health status/disability) / HIC: high-income country.

† syphilis screening, HIV counselling and screening, anaemia screening, pre-clamsia screening, iron and folic Acid use, malaria presumptive intermittent treatment, blood pressure measurement, measurement of symphysis-fundal height, ultrasound, gestational diabetes screening, asymptomatic bacteriuria, Rh type screening

tetanus immunization, calcium use.

Table S1.4. Number of studies by sub-type of **maternal, neonatal and general child-care** interventions, according to T. Tanahashi's coverage steps.

Intervention	Number of interventions assessed	**Availability**	**Accessibility**	**Affordability**	**Acceptability**	**Contact**	**Effective**
Family planning	**3**	1	0	0	0	3	3
Antenatal care (general)	**18**	3	1	1	0	18	17
Antenatal care (18 specific activities) †	**32**	10	10	0	2	32	32
Obstetric care	**2**	1	0	0	1	2	2
Skilled birth attendance	**17**	3	2	0	0	16	15
Postnatal care	**10**	1	1	0	0	10	9
Preventive health control (child)	**1**	0	0	0	0	1	0
Child health services	**2**	0	0	0	0	2	2
Maternal and child health service	**2**	1	0	0	1	2	1
	87	23.0%	16.1%	1.1%	4.6%	98.9%	93.1%

† syphilis screening, HIV counselling and screening, anaemia screening, pre-clamsia screening, iron and folic Acid use, malaria presumptive intermittent treatment, blood pressure measurement, measurement of symphysis-fundal height, ultrasound, gestational diabetes screening, asymptomatic bacteriuria, Rh type screening

tetanus immunization, calcium use.

Table S1.5. Indicator used for utilization and quality in **maternal, neonatal and general child-care** sub-type interventions.

Intervention	**Utilization**	**Quality**
Family planning	Use of a modern contraceptive method	Facility index of quality
Antenatal care (general)	At least 1 antenatal visit/ at least 1 antenatal visit by a skilled health worker/ at least 4 antenatal visits by a skilled health worker	At least 4 visits/ Scores according to essential services*1,*2/ At least 4 visits, first ANC visit before 4 months gestation, score according to essential services*3/ Facility index of quality/ Indirectly measured in terms of incidence of low weight births/ Number of full term delivery
Antenatal care (18 specific activities) †	Multiple	Multiple
Obstetric care	Health care contact	Facility index of quality
Skilled birth attendance	Birth at any place/ Birth attended in a health worker/ Birth attended in a healthcare facility	Birth attended in a health care centre/ Birth attended in a hospital/ Maternal mortality and delivery complications/ Facility index of quality/ Availability of partograph/ Active management of third stage of labour/ Availability of postnatal care and giving an oxytocic agent during the last delivery attended/ Ration between the fraction of caesarean section-WHO standard/ Postnatal check-up within 48 post-delivery at a health care facility
Postnatal care	Delivery in a hospital/ Birth at a heath care facility/ Postnatal visit/ Postpartum check with 48 hours/ Breastfeeding	Activities related to breasts and bleeding, counselling on danger signs, nutrition, family planning/ Weight the new-born, check cord, counsel on breastfeeding, thermal care, discuss danger signs/ Facility index of quality and report of checking during first 7 days after delivery within 48 hours after delivery/ Exclusive breastfeeding/ Incidence of diarrhoea and acute respiratory infection/ Late neonatal mortality due to specific causes/ Mortality rate in premature babies
Preventive health control (child)	At least 1 health control visit during the first year of life	-
Child health services	Sought care at the nearest facility	Facility index of quality
Maternal and child health service	Skilled Antenatal care, plus initial visit during the first trimester of pregnancy, plus at least 4 visits, plus adequate content of antenatal care (8 procedures), plus delivery in healthcare facility, plus + skilled professionals participated in the delivery	Midwife participated in care provision/ No maternal complication during childbirth and normal child weight

*1: blood pressure measured, urine testing, blood sample taken, tetanus toxoid vaccination, first Antenetal care visit before 4 months gestation, iron–folate supplementation, respondent informed about pregnancy complications, drug for intestinal parasites prescribed.
*2: blood pressure measured, urine testing, blood sample taken, tetanus toxoid vaccination, iron–folate supplementation, body weight measured
*3: blood pressure measured, urine testing, blood sample taken, tetanus toxoid vaccination, iron–folate supplementation, body and height weight measured, HIV tcounseling and test offered, iron supplementation provided, antimalarial drug provided for IPT, birth preparedness plan discussed, drug for intestinal parasites prescribed.

Table S1.6. Number of studies by sub-type interventions addressed to **infectious diseases**, that met criteria of utilization according B. Shengelia et al, type of measurement of health gain (quality) and number of countries involved in the assessment of effective coverage.

Intervention	Number of countries were HIC	Number of countries	Number of articles	Number of interventions assessed	**Utilization** is conditioned to a true need	Q1	Q2	Q3	Q4	Q5	Q6
Immunisation **	2	23	**17**	**20**	19	2	0	4	7	0	0
Acute respiratory infections (child)	0	12	**7**	**7**	6	3	1	3	0	0	0
Diarrhoea (child)	0	3	**6**	**6**	5	2	1	1	0	0	0
Lymphatic Filariasis MDA	0	1	**14**	**14**	14	0	0	14	0	0	0
Malaria vector control (ITN/ LLINT/ IRS/ MDA)	0	7	**8**	**9**	9	0	0	6	0	0	0
Malaria treatment	0	46	**6**	**7**	6	0	0	6	1	0	0
Malaria diagnosis	0	1	**1**	**1**	0	-	-	-	-	-	-
HIV counselling and screening	0	2	**2**	**2**	0	2	0	0	0	0	0
HIV treatment	0	0	**2**	**2**	2	0	0	1	0	0	0
Tuberculosis screening	0	1	**1**	**1**	1	-	-	-	-	-	-
Tuberculosis treatment	0	1	**1**	**1**	1	-	-	-	-	-	-
Trachoma MDA	0	1	**1**	**1**	1	0	0	1	0	0	0
	2	51	**38**	**71**	90.1%	12.7%	2.8%	50.7%	11.3%	0.0%	0.0%

Q1: Quality type 1 (health care provider)/ Q2: Quality type 2 (incidence of a disease or condition)/ Q3: Quality type 3 (content of heath care)/ Q4: Quality type 4 (biomarker and function)/ Q5: Quality type 5 (activity/ participation)/ Q6: Quality type 6 (health status/disability) / MDA: massive drug administration/ ITN: insecticide-treated nets/ LLITN: Long-lasting insecticide-treated nets/ IRS: indoor residual spraying/ ACT: Artemisinin – Combination Therapy/ HIC: high-income country.

** immunization include: measles (9); DPT 1 dose -DPT 3 doses (8); BCG (6); OPV (3); the whole immunization programme (3); influenza vaccine (2); tetanus (1); diphtheria (a), rubella (1); mumps (1); Hib (1); Hepatitis B (1)

Table S1.7. Number of studies by sub-type interventions addressed to **infectious diseases**, according to T. Tanahashi's coverage steps.

Intervention	Number of interventions assessed	**Availability**	**Accessibility**	**Affordability**	**Acceptability**	**Contact**	**Effective**
Immunisation **	**20**	1	2	1	2	20	13
Acute respiratory infections (child)	**7**	2	1	1	1	7	7
Diarrhoea (child)	**6**	1	0	0	0	6	4
Lymphatic Filariasis MDA	**14**	1	0	0	14	14	14
Malaria vector control (ITN/ LLINT/ IRS/ MDA)	**9**	2	0	0	2	9	6
Malaria treatment	**7**	1	1	0	1	7	7
Malaria diagnosis	**1**	1	0	0	0	1	0
HIV counselling and screening	**2**	1	0	0	0	1	0
HIV treatment	**2**	0	1	1	1	1	1
Tuberculosis screening	**1**	0	0	0	0	1	0
Tuberculosis treatment	**1**	0	0	0	0	1	0
Trachoma MDA	**1**	0	0	0	0	1	1
	71	14.1%	7.0%	4.2%	29.6%	97.2%	74.6%

MDA: massive drug administration/ ITN: insecticide-treated nets/ LLITN: Long-lasting insecticide-treated nets/ IRS: indoor residual spraying/ ACT: Artemisinin – Combination Therapy

Table S1.8. Indicator used for utilization and quality in sub-type interventions addressed to **infectious diseases**.

Intervention	**Utilization**	**Quality**
Immunisation **	Parental recall/ immunisation card/ administrative records/ presence of a scare in the arm (BCG)	Schedule of immunisation updated/ Seroprevalence/ Facility categorized as high quality
Acute respiratory infections (child)	Health care contact/ mother reports that the child received treatment or advice	Treatment from a health worker/ Did not result in hospitalization/ inputs from literature review/ facility index of quality
Diarrhoea (child)	Health care contact/ mother reports child received more liquids or oral rehydration therapy	Mother receives guideline for rehydration salt mixture/ Did not result in hospitalization/ facility index of quality
Lymphatic Filariasis MDA	The medication is received	Self-report of consumption of the medication/ medication is consumed/ direct observation of medication intake
Malaria vector control (ITN/ LLINT/ IRS/ MDA)	Owner of an ITN (or LLITN)/ ITN (or LLITN) was used/ IRS was used/ the medication is received	ITNs in functionally good condition/ Concertation of insecticide is adequate/ self-report of consumption of the medication/ use of LLITN
Malaria treatment	Treatment was sought/ Health care contact/ the treatment is received/ self-report of treatment	Consume the drug/ Rapid diagnostic tests and prompt treatment / Treatment received within 48 hours (ACT)/ Treatment is completed according to guidelines/ Clinical and parasitological cure
Malaria diagnosis	Self-report of diagnosis by a laboratory test	-
HIV counselling and screening	Health control in the previous year, according to administrative records	-
HIV treatment	Patients receiving antiretroviral therapy according to administrative records	Retention in care
Tuberculosis screening	Ratio between self-report TB-like symptoms-people who had health care contact	-
Tuberculosis treatment	Self-reported having completed the whole treatment protocol	-
Trachoma MDA	The medication is received	The medication is consumed

MDA: massive drug administration/ ITN: insecticide-treated nets/ LLITN: Long-lasting insecticide-treated nets/ IRS: indoor residual spraying/ ACT: Artemisinin – Combination Therapy

Table S1.9. Number of studies by sub-type of interventions addressed to **non-communicable diseases**, that met criteria of utilization according B. Shengelia et al, type of measurement of health gain (quality) and number of countries involved in the assessment of effective coverage.

Intervention	Number of countries were HIC	Number of countries	Number of articles	Number of interventions assessed	**Utilization** is conditioned to a true need	Q1	Q2	Q3	Q4	Q5	Q6
Diabetes treatment	8	12	**11**	**12**	9	0	0	1	11	0	0
Diabetes (6 other preventive activities) ‡	0	1	**1**	**6**	5	0	3	1	2	0	0
Hypertension treatment	2	8	**12**	**12**	10	0	0	1	11	0	0
Hypertension screening	1	0	**1**	**1**	1	0	0	1	0	0	0
Cervical cancer screening	20	63	**5**	**6**	5	0	0	3	0	0	0
Breast cancer screening	1	9	**4**	**4**	3	0	0	2	0	0	0
Hyperlipidaemia treatment	2	4	**3**	**3**	3	0	0	0	3	0	0
Angina treatment	0	1	**1**	**1**	1	-	-	-	-	-	-
Asthma treatment	0	1	**1**	**1**	1	-	-	-	-	-	-
Arthritis treatment	0	1	**1**	**1**	1	-	-	-	-	-	-
Vision disorder treatment	0	1	**1**	**1**	1	0	0	0	0	1	0
	21	64	**18**	**48**	83.3%	0.0%	6.3%	18.8%	56.3%	2.1%	0.0%

Q1: Quality type 1 (health care provider)/ Q2: Quality type 2 (incidence of a disease or condition)/ Q3: Quality type 3 (content of heath care)/ Q4: Quality type 4 (biomarker and function)/ Q5: Quality type 5 (activity/ participation)/ Q6: Quality type 6 (health status/disability) / HIC: high-income country

‡ Screening of diabetes, retinopathy prevention, diabetic foot prevention, nephropathy prevention, hypertension treatment in diabetic, hyperlipidaemia in diabetic

Table S1.10. Number of studies by sub-type interventions addressed to **non-communicable diseases**, according to T. Tanahashi's coverage steps.

Intervention	Number of interventions assessed	**Availability**	**Accessibility**	**Affordability**	**Acceptability**	**Contact**	**Effective**
Diabetes treatment	**12**	1	0	0	1	12	12
Diabetes (6 other preventive activities) ‡	**6**	0	0	0	0	6	6
Hypertension treatment	**12**	0	2	0	2	12	12
Hypertension screening	**1**	0	0	0	0	1	1
Cervical cancer screening	**6**	0	0	0	0	6	3
Breast cancer screening	**4**	0	0	0	0	4	2
Hyperlipidaemia treatment	**3**	0	0	0	0	3	3
Angina treatment	**1**	0	0	0	0	1	0
Asthma treatment	**1**	0	0	0	0	1	0
Arthritis treatment	**1**	0	0	0	0	1	0
Vision disorder treatment	**1**	0	0	0	0	1	1
	48	2.1%	4.2%	0.0%	6.3%	100.0%	83.3%

‡ Screening of diabetes, retinopathy prevention, diabetic foot prevention, nephropathy prevention, hypertension treatment in diabetic, hyperlipidaemia in diabetic

Table S1.11. Indicator used for utilization and quality in sub-type interventions addressed to **non-communicable diseases.**

Intervention	**Utilization**	**Quality**
Diabetes treatment	Self-reported of treatment (medication) / self-report fasting glucose measurement/ self-repot HbA1c measurement/ Diabetics who attended to healthcare provider in the previous year/ diabetic according to administrative records	Reduction in fasting plasma glucose/ HbAc1, compared with treatment targets
Diabetes (6 other preventive activities) ‡	Multiple	Multiple
Hypertension treatment	Self-report of treatment (medication)/ hypertension according to administrative records/ Health control in the previous year	Reduction in blood pressure compared with treatment targets/ Use of health services according to the disease
Hypertension screening	Screening for hypertension	Receives intervention according to level of blood pressure and guidelines
Cervical cancer screening	A pap within past year/ a VPH screening within past year/ a pelvic exam within past year	Women with abnormal results were treated/ timely conformation of diagnosis and treatment/ Pap in the previous 3 years
Breast cancer screening	A mammography within past year	Women with abnormal results were treated/ timely conformation of diagnosis and treatment
Hyperlipidaemia treatment	Self-report of treatment (medication)/ hyperlipidaemia according to administrative records	Reduction in lipidemic levels compared with treatment targets
Angina treatment	Self-report of treatment (medication)	-
Asthma treatment	Self-report of treatment (medication)	-
Arthritis treatment	Self-report of treatment (medication)	-
Vision disorder treatment	Use of glasses or contact lenses	Report no near or far visual impairment when wearing glasses or contact lenses

‡ Screening of diabetes, retinopathy prevention, diabetic foot prevention, nephropathy prevention, hypertension treatment in diabetic, hyperlipidaemia in diabetic

Table S1.12. Number of studies by sub-type of interventions addressed to **mental health diseases**, that met criteria of utilization according B. Shengelia et al, type of measurement of health gain (quality) and number of countries involved in the assessment of effective coverage.

Intervention	Number of countries were HIC	Number of countries	Number of articles	Number of interventions assessed	**Utilization** is conditioned to a true need	Q1	Q2	Q3	Q4	Q5	Q6
Depression	2	3	**3**	**3**	3	0	0	0	1	0	1
Drug and alcohol use disorders	0	3	**3**	**3**	2	0	0	0	2	0	0
Epilepsy	0	3	**2**	**2**	1	0	0	0	2	0	0
Obsessive compulsive disorder	0	1	**1**	**1**	1	0	0	0	1	0	0
Programme	3	5	**5**	**5**	1	0	0	1	0	1	0
Psychosis and severe mental disorders	2	2	**3**	**3**	1	0	0	0	1	1	0
	5	8	**10**	**17**	52.9%	0.0%	0.0%	5.9%	41.2%	11.8%	5.9%

Q1: Quality type 1 (health care provider)/ Q2: Quality type 2 (incidence of a disease or condition)/ Q3: Quality type 3 (content of heath care)/ Q4: Quality type 4 (biomarker and function)/ Q5: Quality type 5 (activity/ participation)/ Q6: Quality type 6 (health status/disability) / HIC: high-income country

Table S1.13. Number of studies by sub-type of interventions addressed to **mental health diseases**, according to T. Tanahashi's coverage steps.

Intervention	Number of interventions assessed	**Availability**	**Accessibility**	**Affordability**	**Acceptability**	**Contact**	**Effective**
Depression	**3**	0	0	0	0	3	2
Drug and alcohol use disorders	**3**	0	0	0	0	3	2
Epilepsy	**2**	0	0	0	0	2	2
Obsessive compulsive disorder	**1**	0	0	0	0	1	1
Programme	**5**	0	0	0	0	5	2
Psychosis and severe mental disorders	**3**	0	0	0	0	3	2
	17	0.0%	0.0%	0.0%	0.0%	100.0%	64.7%

Table S1.14. Indicator used for utilization and quality in sub-type of interventions addressed to **mental health diseases**.

Intervention	**Utilization**	**Quality**
Depression	Self-report of treatment/ in treatment according to administrative records	Disability weights averted from optimal care/ Minimally adequate treatment/ clinical score / score of disability
Drug and alcohol use disorders	Health care contact/ mental health care contact/ in treatment according to administrative records	Abstinent rate/ Retention in treatment/ Minimally adequate treatment/ clinical score / score of disability
Epilepsy	Health care contact/ mental health care contact/ in treatment according to administrative records	Anti-epileptic treatment is received /reduction in number of seizures/ score of disability
Obsessive compulsive disorder	Health care contact/ mental health care contact	Clinical score
Programme	Health care contact/ mental health care contact/ admitted to the programme according to administrative records	Completing course of treatment/ minimally adequate treatment/ recovery
Psychosis and severe mental disorders	Health care contact/ mental health care contact/ admitted to the programme according to administrative records	Adherence to antipsychotic medications/ improved as assessed by the treating clinician/ normal functioning and without risk behaviour such as suicide or violent behaviour/ clinical score / score of disability

Table S1.15. Number of studies by sub-type of **nutritional** interventions, that met criteria of utilization according B. Shengelia et al, type of measurement of health gain (quality) and number of countries involved in the assessment of effective coverage.

Intervention	Number of countries were HIC	Number of countries	Number of articles	Number of interventions assessed	**Utilization** is conditioned to a true need	Q1	Q2	Q3	Q4	Q5	Q6
Nutrition Prevention	0	7	**14**	**17**	17	0	0	9	8	0	0
Nutrition Treatment	0	1	**1**	**1**	1	0	0	0	1	0	0
	0	8	**15**	**18**	100.0%	0.0%	0.0%	50.0%	50.0%	0.0%	0.0%

Q1: Quality type 1 (health care provider)/ Q2: Quality type 2 (incidence of a disease or condition)/ Q3: Quality type 3 (content of heath care)/ Q4: Quality type 4 (biomarker and function)/ Q5: Quality type 5 (activity/ participation)/ Q6: Quality type 6 (health status/disability) / HIC: high-income country

Table S1.16. Number of studies by sub-type **nutritional** interventions, according to T. Tanahashi's coverage steps.

Intervention	Number of interventions assessed	**Availability**	**Accessibility**	**Affordability**	**Acceptability**	**Contact**	**Effective**
Nutrition Prevention	**17**	0	0	1	17	17	0
Nutrition Treatment	**1**	0	0	0	1	1	0
	18	0.0%	0.0%	5.6%	100.0%	100.0%	0,0%

Table S1.17. Indicator used for utilization and quality in sub-type of **nutritiona**l interventions.

Intervention	**Utilization**	**Quality**
Nutrition Prevention	The child has ever been fed the product/ Consumes the supplement sometimes or always/ Consumes the vehicle.	The child has ever been fed the product in accordance to programme goals/ Achieves sufficiency intake of the nutrient/ Consumes the fortified vehicle/ Consumes the fortified vehicle adequately/ Consumes the supplement always.
Nutrition Treatment	Admission to malnutrition.	Cured from malnutrition service.

Table S1.18. Number of studies by sub-type of interventions addressed to **injuries and violence**, that met criteria of utilization according B. Shengelia et al, type of measurement of health gain (quality) and number of countries involved in the assessment of effective coverage.

Intervention	Number of countries were HIC	Number of countries	Number of articles	Number of interventions assessed	**Utilization** is conditioned to a true need	Q1	Q2	Q3	Q4	Q5	Q6
Domestic violence against women	0	1	**1**	**1**	1	0	0	1	0	0	0
	0	1	**1**	**1**	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%

Q1: Quality type 1 (health care provider)/ Q2: Quality type 2 (incidence of a disease or condition)/ Q3: Quality type 3 (content of heath care)/ Q4: Quality type 4 (biomarker and function)/ Q5: Quality type 5 (activity/ participation)/ Q6: Quality type 6 (health status/disability) / HIC: high-income country

Table S1.19. Number of studies by sub-type of interventions addressed to **injuries and violence**, according to T. Tanahashi's coverage steps.

Intervention	Number of interventions assessed	**Availability**	**Accessibility**	**Affordability**	**Acceptability**	**Contact**	**Effectiveness**
Domestic violence against women	**1**	0	0	0	0	1	1
	1	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%

Table S1.20. Indicator used for utilization and quality in sub-type of interventions addressed to **injuries and violence.**

Intervention	**Utilization**	**Quality**
Domestic violence against women	Self-reported use of health services associated with manifestations of violence	The staff recommended to report the perpetrator to the police authorities

Table S1.21. Number of studies by sub-type of interventions addressed to **sanitation and others**, that met criteria of utilization according B. Shengelia et al, type of measurement of health gain (quality) and number of countries involved in the assessment of effective coverage.

Intervention	Number of countries were HIC	Number of countries	Number of articles	Number of interventions assessed	**Utilization** is conditioned to a true need	Q1	Q2	Q3	Q4	Q5	Q6
Smoking cessation	0	1	**1**	**1**	1	-	-	-	-	-	0
Sanitary toilet	0	1	**2**	**2**	2	0	0	1	0	0	0
Drinking water	0	1	**1**	**1**	1	-	-	-	-	-	-
	0	1	**2**	**4**	100.0%	0.0%	0.0%	25.0%	0.0%	0.0%	0.0%

Q1: Quality type 1 (health care provider)/ Q2: Quality type 2 (incidence of a disease or condition)/ Q3: Quality type 3 (content of heath care)/ Q4: Quality type 4 (biomarker and function)/ Q5: Quality type 5 (activity/ participation)/ Q6: Quality type 6 (health status/disability) / HIC: high-income country

Table S1.22. Number of studies by sub-type of interventions addressed to **sanitation and others**, according to T. Tanahashi's coverage steps.

Intervention	Number of interventions assessed	**Availability**	**Accesibility**	**Affordability**	**Acceptability**	**Contact**	**Effectiveness**
Smoking cessation	**1**	0	0	0	0	1	0
Sanitary toilet	**2**	0	0	0	0	2	1
Drinking water	**1**	0	0	0	0	1	0
	4	0.0%	0.0%	0.0%	0.0%	100.0%	25.0%

Table S1.23. Indicator used for utilization and quality in sub-type of interventions addressed to **sanitation and others.**

Intervention	**Utilization**	**Quality**
Smoking cessation	Smokers who either quit smoking or tried to quit	-
Sanitary toilet	Households with access to a sanitary toilet/ Availability of shared sanitation	Compliance in the use of shared sanitation
Drinking water	Households with access to safe drinking water	-

Table S1.24

Id	Authors	Year	Countries	Target interventions	Type of intervention	Utilization conditioned to a true need?	Quality type	Study design	Indicator of utilization	Indicator of quality
1	Cesar et al.[55]	1986	Brazil	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	-	Number of antenatal care visits
2	al-Sekait et al.[56]	1989	Saudi Arabia	Maternal and child health service	Maternal and child health service	Yes		Cross-sectional	-	If midwife participated in care provision
3	Andrews et al.[44]	2000	Australia	Depression	Mental health: Depression	Yes	Health status/disability	Cost-effectiveness study	-	Disability weights averted from optimal care
4	Borus [57]	2004	Kenya	BCG, OPV, DPT	Immunisation	Yes	Content care	Cross-sectional	-	To keep the schedule updated of number of doses
5	Lozano et al. [43]	2006	México	BCG, DPT, Measles	Immunisation	Yes	-	Cross-sectional	(BCG: scare/ DPT: at least 3 doses/ Measles: 1 or more doses)	-
				Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 4 antenatal visits (doctor, nurse, or midwife)	Received blood test and had blood pressure measured
				Skilled birth attendance	Skilled birth attendance	Yes	Content care	Cross-sectional	Birth attended by health professional	Birth took place in hospital
				Care of premature neonates	Postnatal care	Yes	Incidence of disease	Cross-sectional	Birth took place in hospital	Difference in mortality rate in premature babies compared with maximum and minimum risk-adjusted mortality
				Diarrhoea	Diarrhoea	Yes	-	Cross-sectional	Mother reports child received more liquids or oral rehydration therapy	-
				Acute respiratory infections	Acute respiratory infections	Yes	Content care	Cross-sectional	Mother reports child received treatment	Treatment from a health worker
				Breast cancer screening	Breast cancer screening	Yes	-	Cross-sectional	Had a mammography within past year	-

				Cervical cancer screening	Cervical cancer screening	Yes	-	Cross-sectional	Had a pap within the last year	-
				Treatment of vision disorders	Vision disorder treatment	Yes	Activity	Cross-sectional	Use glasses or contact lenses	Report no near or far visual impairment when wearing glasses or contact lenses
				Treatment of asthma	Asthma treatment	Yes	-	Cross-sectional	Self-reported medication	-
				Treatment of angina	Angina treatment	Yes	-	Cross-sectional	Self-reported medication	-
				Treatment of arthritis	Arthritis treatment	Yes	-	Cross-sectional	Self-reported medication	-
				Treatment of diabetes	Diabetes treatment	Yes	Biomarker of funcionality	Cross-sectional	Self-reported use of oral hypoglycaemics or insulin	Reduction in fasting plasma glucose compared with treatment targets
				Treatment of hypertension	Hypertension treatment	Yes	Biomarker of funcionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets
				Treatment of hypercholesterolaemia	Dyslipidaemia treatment	Yes	Biomarker of funcionality	Cross-sectional	Self-reported use of drugs for cholesterol reduction	Reduction in total cholesterol compared with treatment targets
				Influenza vaccine	Immunisation	Yes	-	Cross-sectional	Self-reported influenza vaccine in past year	-
6	Gakidou et al. [58]	2006	Same information than Lozano et al 2006	-		-		-	-	-
7	Gakidou et al. [59]	2008	52 countries	Cervical cancer screening	Cervical cancer screening	Yes	Content care	Cross-sectional	women who report that they have had a pelvic exam	pelvic exam and Pap smear in the past three years
8	Kumar et al. [60]	2008	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
9	Liu et al. [47]	2008	China	Access to safe drinking water	Drinking water	Yes	-	Cross-sectional	Households with access to safe drinking water	-
				Access to sanitary toilets	Sanitary toilet	Yes	-	Cross-sectional	Households with access to a sanitary toilet	-
				Smoking cessation	Smoking cessation	Yes	-	Cross-sectional	Smokers who either quit smoking or tried to quit in the year before the survey	-
				Antenatal care	Antenatal care	Yes	-	Cross-sectional	At least 1 antenatal visit	-

				Skilled birth attendance	Skilled birth attendance	Yes	-	Cross-sectional	Women who gave birth, delivering their baby in a hospital in the year before the survey	-
				Postnatal care	Postnatal care	Yes	-	Cross-sectional	Women who gave birth in the year before the survey, who received any postnatal visit by medical staff	-
				BCG, DTP3, measles, and Hep B	Immunisation	Yes	-	Cross-sectional	Children younger than 1 year who were immunised for all of the agents	-
				Examination of suspected TB cases	TB screening	Yes	-	Cross-sectional	Ratio of people who reported having gone through formal clinical examination to people who reported having TB-like symptoms	-
				Treatment of confirmed TB cases	TB treatment	Yes	-	Cross-sectional	Confirmed TB patients who self-reported having completed the whole treatment protocol	-
				Treatment of hypertension	Hypertension treatment	Yes	Biomarker of functionality	Cross-sectional	Hypertensive people who reported having taken control measures in the past year	Reduction in systolic blood pressure compared with treatment targets
10	Hussein at al. (a) [61]	2011	Ethiopia	HIV test counselling and HIV test	HIV counselling and screening	No	Health provider quality	Cross-sectional	Number of appointments according to records	Quality was measured at centre level, while results of utilization are described at the individual level
11	Hussein at al. (b) [62]	2011	Ethiopia	HIV test counselling, HIV test, counselling on infant feeding options and women who took Nevirapine	HIV counselling and screening	No	Health provider quality	Cross-sectional	Number of appointments according to records	Quality was measured at centre level, while results of utilization are described at the individual level
12	Idzerda et al. [63]	2011	Serbia	Acute respiratory infections	Acute respiratory infections	Yes	Content care	Cross-sectional	Health care contact	Combination of efficacy of antibiotic and diagnostic accuracy using inputs from a literature review
13	Martinez et al. [64]	2011	Bolivia, Colombia,	BCG, DTP3, VOP, measles	Immunisation	Yes	-	Cross-sectional	(BCG: at least 1 dose/ DPT: at least 3 doses/	-

			Chile, Costa Rica, Haití, Honduras, México, Dominican Republic, Peru						VOP: at least 3 doses/ Measles: at least doses)	
				Healthy child control	Health control of child	Yes	-	Cross-sectional	At least 1 control	-
				Breastfeeding	Postnatal care	Yes	Incidence of disease	Cross-sectional	Mother self-report of breastfeeding	Indirectly measured in terms of incidence of diarrhoea and ARI
				Antenatal care	Antenatal care	Yes	Incidence of disease	Cross-sectional	At least 4 antenatal visits (doctor, nurse, or midwife)	Indirectly measured in terms of incidence of low weight births
				Skilled birth attendance	Skilled birth attendance	Yes	Incidence of disease	Cross-sectional	Birth attended by health professional	Indirectly measured in terms of maternal mortality and delivery complications
				Institutionalised birth attendance	Skilled birth attendance	Yes	-	Cross-sectional	Birth attended in a healthcare facility	-
				Breast cancer screening	Breast cancer screening	Yes	Content care	Cross-sectional	Had a mammography within past year	Women with abnormal results were treated
				Cervical cancer screening	Cervical cancer screening	Yes	Content care	Cross-sectional	Had a pap within the last year	Women with abnormal results were treated
				Treatment of diabetes	Diabetes treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of oral hypoglycaemics or insulin	Reduction in fasting plasma glucose compared with treatment targets
				Treatment of hypertension	Hypertension treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets
				Treatment of hypercholesterolaemia	Dyslipidaemia treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported treatment for cholesterol reduction	Reduction in total cholesterol compared with treatment targets
14	Lopez-Lopez et al. [65]	2012	México	Diabetes detection	Diabetes screening	No	Content care	Cross-sectional	Number in the programme	Assistance on two or more occasions
				Treatment of diabetes	Diabetes treatment	Yes	Biomarker of functionality	Cross-sectional	Fasting glucose measurement	Reduction in fasting plasma glucose compared with treatment targets
				Treatment of diabetes	Diabetes treatment	Yes	Biomarker of functionality	Cross-sectional	HbA1c measurement	Reduction in HbA1c compared with treatment targets
				Treatment of diabetes with hyperlipidaemia	Diabetes dyslipidaemia treatment	Yes	Biomarker of functionality	Cross-sectional	Lipidemia measurement	Reduction in lipidemia compared with treatment targets
				Treatment of diabetes with hypertension	Diabetes hypertension treatment	Yes	Biomarker of functionality	Cross-sectional	Blood pressure measurement	Reduction in systolic blood pressure

										compared with treatment targets
				Prevention of diabetic retinopathy	Diabetic prevention retinopathy	Yes	Disease incidence	Cross-sectional	Evaluation with ophthalmoscope	Without retinopathy
				Prevention of diabetic foot	Diabetic prevention foot	Yes	Disease incidence	Cross-sectional	2 or more evaluations per year	Without diabetic foot
				Prevention of diabetic nephropathy	Diabetic prevention nephropathy	Yes	Disease incidence	Cross-sectional	Annual measurement	Without nephropathy
15	Shreyash et al [66]	2012	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
16	Ranganath et al. [67]	2012	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
17	Knaul et al. [68]	2012	Same information than Gutierrez 2013, et al. and Lozano et al 2006	-	-	-	-	-	-	-
18	Beer et al. [69]	2013	Tanzania	Malaria vector control interventions in children (insecticide-treated nets (ITNs) and indoor residual spraying (IRS))	Malaria vector control (ITN + IRS)	Yes	-	Cross-sectional	-	It is assumed that Vector control intervention are "effective"
19	Ghosh et al. [70]	2013	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
20	Gutierrez [71]	2013	México	BCG, DPT, Measles	Immunisation	Yes	-	Cross-sectional	(BCG: scare/ DPT: at least 3 doses/ Measles: 1 or more doses)	-
				Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 4 antenatal visits (doctor, nurse, or midwife)	Received blood test and had blood pressure measured
				Skilled birth attendance	Skilled birth attendance	Yes	Content care	Cross-sectional	Birth attended by health professional	Birth took place in hospital
				Diarrhoea	Diarrhoea	Yes	-	Cross-sectional	Mother reports child received more liquids or oral rehydration therapy	-
				Acute respiratory infections	Acute respiratory infections	Yes	Content care	Cross-sectional	Mother reports that the child received treatment	Treatment from a health worker
				Breast cancer screening	Breast cancer screening	Yes	-	Cross-sectional	Had a mammography within past year	-
				Cervical cancer screening	Cervical cancer screening	Yes	-	Cross-sectional	Had a pap within the last year	-
				Influenza vaccine	Immunisation	Yes	-	Cross-sectional	Self-reported influenza vaccine in past year	-

				VPH screening	Cervical cancer screening	Yes	-	Cross-sectional	Had a VPH screening within the last year	-
21	Nesbitt et al. [72]	2013	Ghana	Skilled birth attendance	Skilled birth attendance	Yes	Health provider quality	Cross-sectional	Birth attended in a health facility	Facility categorised as high quality
22	Roy et al. [73]	2013	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug (optimal dose)
23	Scheibe et al. [74]	2013	Uganda	HIV antiretroviral therapy	HIV treatment antiretroviral	Yes	Content care	Cross-sectional	Received the drug	Received and retention in care
24	Hitesh et al. [75]	2013	India	DTP first dose	Immunisation	Yes	Health provider quality	Cross-sectional	DPT 1 dose	Facility categorised as high quality
25	Singh et al.	2013	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
26	Viviescas-Vargas et al. [76]	2013	México	Health services for domestic violence against women	Domestic Violence Women	-	Content care	Cross-sectional	Self-reported use of health services associated with manifestations of violence	The staff recommended to report the perpetrator to the police authorities in accordance with country regulations
27	Adhikari et al. [77]	2014	Nepal	Malaria diagnosis	Malaria diagnosis	No	-	Cross-sectional	Self-report of diagnosis by a laboratory test	-
				Malaria treatment	Malaria treatment	Not clear	Content care	Cross-sectional	Self-reported treatment	Treatment according to national guidelines
				Malaria vector control: LLIN (long-lasting insecticide-treated nets) intervention	Malaria vector control (ITN)	Yes	-	Cross-sectional	-	It is assumed that Vector control intervention are "effective"
				Malaria vector control: IRS (indoor residual spraying) intervention	Malaria vector control (IRS)	Yes	-	Cross-sectional	-	It is assumed that Vector control intervention are "effective"
28	De Sylva [41]	2014	-	-		-		Systematic Review	-	-
29	Department of Health [45]	2012	England	Mental Health Programme for common mental disorders	Mental health: Programme	No	activity/ participation	Cross-sectional	Information from records from health care	Completing course of treatment/ recovery
30	Pirkis et al. [78]	2011	Australia	Mental Health Programme for any mental disorder	Mental health: Programme	No	-	Cross-sectional	Information from records from health care	-
31	Araya et al. [79]	2018	Chile	Mental Health Programme for depression	Mental health: Depression	Yes	-	Cross-sectional	-	-
32	Aagaard et al. [80]	2000	Denmark	Mental Health Programme for Severe Mental Illness	Mental health: Psychosis and severe mental disorders	No	-	Cross-sectional	Information from records from health care	-

33	Lin et al. [81]	2010	China	Mental Health Programme for Drug use disorder	Mental health: Drug and substance use disorder	No	-	Cross-sectional	Information from records from health care	-
34	Martini et al. [82]	1985	Italia	Mental Health Programme for any mental disorder	Mental health: Programme	No	-	Cross-sectional	Information from records from health care	-
35	Marinoni et al. [83]	1983	Italia	Mental Health Programme for any mental disorder	Mental health: Programme	No	-	Cross-sectional	Information from records from health care	-
36	Hodgins [84]	2014	42 countries	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit	Recived blood pressure measurement, tetanus toxoid vaccination, first ANC visit before 4 months gestation, urine testing, counseling about pregnancy danger signs, and iron–folate supplementation
37	Praveen et al. [85]	2014	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
38	Srivastava et al. [86]	2014	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
39	Zhou et al. [87]	2014	Kenya	Malaria vector control: ITN (insecticide-treated nets) intervention	Malaria vector control (ITN)	Yes	Content care	Cross-sectional	ITN is used	ITNs in functionally good condition
40	Baker et al. (a) [88]	2015	Tanzania, Uganda	Antenatal care Syphilis screening	Antenatal care syphilis screening	Yes	Content care	Cross-sectional	Attending at least once	Received the result
				Antenatal care HIV screening	Antenatal care HIV counselling and screening	Yes	Content care	Cross-sectional	Attending at least once	Received the result
				Antenatal care Anaemia screening	Antenatal care anaemia screening	Yes	Content care	Cross-sectional	Attending at least twice	Leave blood
41	Baker et al. (b) [89]	2015	Tanzania	Antenatal care Syphilis screening	Antenatal care syphilis screening	Yes	Content care	Cross-sectional	Attending at least once	Report having a blood test and receiving a test result
				Antenatal care Pre-eclampsia screening	Antenatal care pre-clamsia screening	Yes	Content care	Cross-sectional	Attending at least three times	Report having their blood pressure checked
				Skilled birth attendance: Use of partograph	Skilled birth attendance	Yes	Health provider quality	Cross-sectional	Birth attended in a health facility	Use a facility with availability of a partograph
				Skilled birth attendance: Active management of the third stage of labour	Skilled birth attendance	Yes	Health provider quality	Cross-sectional	Birth attended in a health facility	Use a facility with availability of postnatal care + giving an oxytocic agent during the last delivery attended

				Postnatal care	Postnatal care	Yes	Health provider quality	Cross-sectional	Birth attended in a health facility	Use a facility with availability of postnatal care + report being checked within 48 hours of delivery
42	Colson et al. [90]	2015	Mexico, Nicaragua	Measles immunisation		Yes		Cross-sectional	-	Seroprevalence
43	Crowe et al. [91]	2015	United Kingdom	Immunisation Programme	Immunisation	Yes	biomarker of funcionality	Simulation methods	-	Seroprevalence
44	Engle-Stone et al. [92]	2015	Cameroon	Vitamin A programme (periodic, VA capsules, Deworming tables, refined oil fortification, bouillon cube fortification, micronutrient powder, biofortified maize)	Nutrition prevention	Yes	functionality	Cost-effectiveness analysis	-	Achieves sufficiency Vitamin A intake
45	Galactionova et al. [93]	2015	43 Sub-Saharan Countries	Malaria treatment	Malaria treatment	Yes	functionality	Cross-sectional	Received the drug	Clinical and parasitological cure
46	Marathe et al. [94]	2015	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
47	Mishra et al. [95]	2015	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
48	Aaron et al. [96]	2016	Ghana	Child Complementary Feeding Supplement	Nutrition prevention	Yes	Content care	Cross-sectional	Child ever been fed the product	Child ever been fed the product in the previous 7 days
49	Luo et al. [97]	2016	Cameroon	Folate in woman and young children	Nutrition prevention	Yes	functionality	Cross-sectional	-	Achieves sufficiency Folate intake
				Vitamin B12 in woman and young children	Nutrition prevention	Yes	functionality	Cross-sectional	-	Achieves sufficiency Vitamin B12 intake
50	Koulidiati et al. [98]	2016	Burkina Faso	BCG, OPV, DPT, Measles (whole programme)	Immunisation	Yes	-	Cross-sectional	Complete the whole programme	-
				Child health services	Child health services	Yes	Health provider quality	Cross-sectional	-	Availability of relevant drugs, equipment and the staffing.
51	Leyvraz et al. [99]	2016	Cote d'Ivoire	Fortified complementary food	Nutrition prevention	Yes	Content care	Cross-sectional	Child ever been fed the product	Child ever been fed the product in the previous 7 days
52	Nguyen et al. [100]	2016	Vietnam	Fortified complementary food	Nutrition prevention	Yes	Content care	Cross-sectional	Child ever been fed the product	Child ever been fed the product in the previous 7 days with at least 3 sachets

53	Patel et al. [101]	2016	China, India	Mental Health Programme for any mental disorder	Mental health: Programme	Yes	Content care	Cross-sectional	Patients having contact with any health-care providers/ Mental health-care provider	(≥1 month of medication and ≥four visits to a medical doctor) or psychotherapy (≥eight visits with a professional) =1-year rate
				Mental Health Programme for severe Mental Illness	Mental health: Psychosis and severe mental disorders	Yes	activity/ participation	Cross-sectional	Patients having contact with any health-care providers/ Mental health-care provider	Estimated adherence to antipsychotic medications according to doses, in past month/ Marked improvement defined as IDEAS score reduction/ improved as assessed by the treating clinician/ normal functioning and without risk behaviour such as suicide or violent behaviour
				Mental Health Programme for Obsessive Compulsive Disorder	Mental health: Obsessive compulsive disorder	Yes	functionality	Cross-sectional	Patients having contact with any health-care providers/ Mental health-care provider	Remission on YBOCS
				Mental Health Programme for Substance Use Disorder	Mental health: Drug and substance use disorder	Yes	functionality	Cross-sectional	Patients having contact with any health-care providers/ Mental health-care provider	Abstinent rate/ Retention in treatment/
				Mental Health Programme Epilepsy	Mental health: Epilepsy	Yes	functionality	Cross-sectional	Patients having contact with any health-care providers/ Mental health-care provider	Responders (≥ 50% reduction in seizure frequency after treatment)/ Rate of receiving any anti-epileptic treatment
54	Bivol et al. [102]	2016	Moldova	HIV antiretroviral therapy	HIV treatment antiretroviral	Yes	-	Cross-sectional	Received the drug	-
55	Servan-Mori et al. 2016[103]	2016	-	Same information than Gutierrez 2013, et al and Lozano et al 2006		-		-	-	-
56	Yawson et al. [104]	2016	Ghana	Postnatal care	Postnatal care	Yes	Health provider quality	Cross-sectional	Birth attended in a health facility	Use a facility with availability of postnatal care + report receiving at least 2 visits in the previous 7 days

57	Comfort et al. [105]	2017	Magadascar	Malaria vector control: ITN (insecticide-treated nets) intervention	Malaria vector control (ITN)	Yes	Content care	RCT	Owner of an ITN	Use an ITN
58	Engle-Stone et al. [106]	2017	Cameroon	Vitamin A programme (periodic, VA capsules, Deworming tables, refined oil fortification, bouillon cube fortification, micronutrient powder, biofortified maize)	Nutrition prevention	Yes	functionality	Cost-effectiveness analysis	-	Achieves sufficiency Vitamin A intake
59	Guerrero-Nunez et al. [50]	2017	Chile	Treatment of diabetes	Diabetes treatment	No	Biomarker of funcionality	Cross-sectional	Not clear	Reduction HbA1c compared with treatment targets
60	Luo et al. [107]	2017	Cameroon	Vitamin A fortification	Nutrition prevention	Yes	functionality	Cross-sectional	-	Achieves sufficiency Vitamin A intake
				Iron fortification	Nutrition prevention	Yes	functionality	Cross-sectional	-	Achieves sufficiency Iron
61	Khan et al. [108]	2017	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
62	Kiwanuka et al. [109]	2017	Uganda	Antenatal care use of iron and folic acid	Antenatal care Iron and Folic Acid use	Yes	Content care	Cross-sectional	Took tablets or syrup	Took tablets or syrup for 90 days.
				Antenatal care intermittent presumptive treatment for malaria	Antenatal care malaria presumptive intermittent treatment	Yes	Content care	Cross-sectional	Attended ANC clinic during the pregnancy.	Received at least two doses of malaria prevention medicine
				Antenatal care HIV counselling and test	Antenatal care HIV counselling and screening	Yes	Content care	Cross-sectional	Got information related to HIV/AIDS and on being tested for the HIV virus.	Received HIV test results.
				Antenatal care Syphilis screening	Antenatal care syphilis screening	Yes	Content care	Cross-sectional	Attended ANC clinic during the pregnancy.	Received syphilis test results
63	Larson et al. [110]	2017	Tanzania	Obstetric care	Obstetric care	Yes	Health provider quality	Cross-sectional	-	Use a facility index of quality
64	Leslie et al. [111]	2017	Haiti, Kenya, Malawi, Namibia, Rwanda, Senegal, Tanzania, Uganda	Antenatal care	Antenatal care	Yes	Health provider quality	Cross-sectional	At least 1 antenatal visit	Use a facility index of quality
				Family planning	Family planning	Yes	Health provider quality	Cross-sectional	Use of a modern contraceptive method	Use a facility index of quality
				Diarrhoea (child)	Diarrhoea	Yes	Health provider quality	Cross-sectional	Health care contact	Use a facility index of quality
				Acute respiratory infections (child)	Acute respiratory infections	Yes	Health provider quality	Cross-sectional	Health care contact	Use a facility index of quality

65	Leyvraz et al. [112]	2017	Ghana, Cote d'Ivoire, India, Bangladesh, Vietnam	Fortified complementary foods	Nutrition prevention	Yes	Content care	Cross-sectional	Child ever been fed the product	Child ever been fed the product according to programme goals
66	Lindtjorn et al. [113]	2017	Ethiopia	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit	Number of visits/ facility quality
				Skilled birth attendance	Skilled birth attendance	Yes	Health provider quality	Cross-sectional	-	Delivery at heath care centre/ facility quality
67	Nguhiu et al. [114]	2017	Kenya	Family planning	Family planning	Yes	Health provider quality	Cross-sectional	Use of a modern contraceptive method	Use a facility index of quality
				Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit	Score according to blood pressure taken, urine sample taken, blood sample taken, respondent informed about pregnancy complications, iron tablets/syrup prescribed, and a drug for intestinal parasites prescribed.
				Skilled birth attendance	Skilled birth attendance	Yes	Health provider quality	Cross-sectional	Birth attended by health professional	Use a facility index of quality
				Breastfeeding first 6mo	Postnatal care	Yes	Content care	Cross-sectional	Mother self-report of breastfeeding during last 24 hours	Exclusive breastfeeding
				Inmunisation Programme	Inmunisation	Yes	Health provider quality	Cross-sectional	Complete the whole programme	Use a facility index of quality
				Diarrhoea (child)	Diarrhoea	Yes	Content care	Cross-sectional	Mother reports child received more liquids or oral rehydration therapy	Receive guideline recommended oral rehydration salt mixture
				Acute respiratory infections (child)	Acute respiratory infections	Yes	Health provider quality	Cross-sectional	Mother reports that the child received advice or treatment	Use a facility index of quality
				Malaria vector control: ITN (insecticide-treated nets) intervention	Malaria vector control (ITN)	Yes	Content care	Cross-sectional	Owner of an ITN	Use an ITN
68	Gabert et al. [115]	2017	India	Treatment of diabetes	Diabetes treatment	Yes	Biomarker of funcionality	Cross-sectional	Self-reported use of oral hypoglycaemics or insulin or fasting blood glucose ≥126 mg/dl or random blood glucose ≥200 mg/dl	Reduction in fasting plasma glucose compared with treatment targets

				Treatment of hypertension	Hypertension treatment	Yes	Biomarker of funcionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets
69	Rios-Blancas et al. [116]	2017	Mexico	Treatment of hypertension	Hypertension treatment	Yes	Biomarker of funcionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets
70	Vosti et al. [117]	2017	Cameroon	Vitamin A fortification (reproductive age women)	Nutrition prevention	Yes	functionality	Cross-sectional	-	Achieves sufficiency Vitamin A intake
				Acid Folic fortification (reproductive age women)	Nutrition prevention	Yes	functionality	Cross-sectional	-	Achieves sufficiency acid folic intake
71	Sharma et al. [118]	2017	Kenya	Antenatal care	Antenatal care	No	Health provider quality	Cross-sectional	Not clear	Use a facility index of quality
				Skilled birth attendance	Skilled birth attendance	No	Health provider quality	Cross-sectional	Not clear	Use a facility index of quality
72	Yawson et al. [104]	2017	Ghana	Immunisation Programme	Immunisation	No	Content care	Cross-sectional	Received first pentavalent vaccine in the previous year (Children <12m)	Complete the whole programme in the previous year (Children <12m)
73	Arredondo et al. [119]	2018	Mexico	Treatment of hypertension	Hypertension treatment	No	Content care	Cross-sectional	Not clear	Use of health services
				Treatment of diabetes	Diabetes treatment	No	Content care	Cross-sectional	Not clear	Use of health services
74	Arsenault et al. [120]	2018	91 LMIC	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit	Receive receipt of three essential services (blood pressure monitoring and urine and blood testing)
75	Banerjee et al. [121]	2018	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
76	Bhatia et al. [122]	2018	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug (optimal dose)
77	Brenner et al. [123]	2018	Malawi	Obstetric care	Obstetric care	Yes	Health provider quality	Cross-sectional	-	Use a facility index of quality
78	Carter et al. [124]	2018	Zambia	Diarrhoea (child)	Diarrhoea	Yes	Health provider quality	Cross-sectional	Health care contact	Use a facility index of quality
				Acute respiratory infections (child)	Acute respiratory infections	Yes	Health provider quality	Cross-sectional	Health care contact	Use a facility index of quality
79	Charoendee et al. [125]	2018	Thailand	Screening of hypertension	Hypertension screening	Yes	Content care	Cross-sectional	Screened for hypertension	Receive intervention according to level of blood pressure and guidelines (risk assessment, disease

										detection, monitoring blood pressure, prevention of CV complications)
80	Jannati et al. [126]	2018	-	-		-		Systematic Review	-	-
81	Randive et al. [127]	2013	India	Skilled birth attendance: Institutional delivery	Skilled birth attendance	No	Health provider quality	Cross-sectional	Proportion of deliveries in an healthcare institution	As a proxy the proportion of caesarean section/standard WHO (5%)
82	Travassos et al. [128]	2016	Ethiopia	DPT (tetanus), Hib	Immunisation	Yes	biomarker of funcionality	Cross-sectional	Parental recall/ immunisation card/ records	Seroprevalence
83	Colson et al. [129]	2013	México	Measles	Immunisation	Yes	biomarker of funcionality	Cross-sectional	Parental recall/ immunisation card	Seroprevalence
84	Just et al. [48]	2018	-	Diarrhoea: shared sanitation facilities	Sanitary toilet	Yes	Content care	Simulation methods	Availability of shared sanitation (coverage)	Use of shared sanitation (compliance)
85	Koulidiati et al. [130]	2018	Burkina Faso	Morbidity healthcare children	Child health services	Yes	Health provider quality	Cross-sectional	Sought care at the nearest facility	Use a facility index of quality
86	Leyvraz et al. [131]	2018	Kenya	Fortified complementary foods (micronutrient powders) for children	Nutrition prevention	Yes	Content care	Cross-sectional	Had ever received MNP for their child	Child had been given the MNP at least 3 times in the previous 7 days
87	Mahdavi et al. [132]	2018	Finland, Germany, Greece, Netherland, Spain, United Kingdom	Treatment of diabetes	Diabetes treatment	Yes	Biomarker of functionality	Cross-sectional	Not clear	Reduction in HbAc1 compared with treatment targets
88	Murphy et al. [133]	2018	Kenya	Small and sick new-born care	Postnatal care	Yes	Health provider quality	Cross-sectional	-	Use a facility index of quality
89	Stones et al. [134]	2018	Malawi	Antenatal care Blood pressure measurement	Antenatal care blood pressure measurement	Yes	Content care	Cross-sectional	Antenatal visit	Blood pressure measurement
90	Willey et al. [135]	2018	Uganda	Skilled birth attendance	Skilled birth attendance	Yes	Health provider quality	Cross-sectional	Birth at any place	Use a facility index of quality
91	Amouzou et al. [42]	2019	-	RMNCH+N	Maternal and child health service and Nutrition	-		Systematic review	-	-
92	Thapa et al. [136]	2016	Nepal	Antenatal care calcium use	Antenatal care calcium use	Yes	Content care	Cross-sectional	Provided with calcium supplementation	Full compliance of calcium supplementation consumption
93	Deming et al. [137]	2002	Central African Republic	Antenatal care tetanus toxoid vaccination	Antenatal care tetanus toxoid vaccination	Yes	biomarker of functionality	Cross-sectional	Self-report of tetanus toxoid vaccination	Seroprevalence

94	Aaron et al. (b) [138]	2016	India	Fortified complementary food	Nutrition prevention	Yes	Content care	Cross-sectional	Consumption of the vehicle	Consumption adequately of the fortified vehicle
95	Aaron et al. [139]	2017	Cote de Ivoire, India, Nigeria, Senegal, South Africa, Tanzania, Uganda, Bangladesh	Fortified complementary food	Nutrition prevention	Yes	Content care	Cross-sectional	Consumption of the vehicle	Consumption of the fortified vehicle
96	Agha et al. [140]	2016	Pakistan	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit	Received blood pressure measurement, blood sample testing, urine sample testing, body weight measurement, whether they were given iron tablets and whether they received tetanus immunizations.
97	Gebremedhin et al. [141]	2014	Ethiopia	Antenatal care use of iron and folic acid	Antenatal care Iron and Folic Acid use	Yes	Content care	Cross-sectional	Given/prescribed iron supplements during pregnancy	Took supplements for >90 consecutive days
98	Hayford et al. [142]	2013	Bangladesh	Measles immunisation	Immunisation	Yes	Biomarker of functionality	Cross-sectional	Parent recall/ card	Seroprevalence
99	Heredia-PI et al. [143]	2016	Mexico	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	Timely ANC (ANC in the first trimester of pregnancy), sufficient ANC (at least 4 ANC visits), appropriate in content (summary of 8 procedures: weight, height, blood pressure, urine analysis, blood analysis, tetanus, vaccination, prescription of folic acid, prescription of vitamin iron or dietary supplements)
100	Kanyangarara et al. [144]	2017	13 sub-Sahara countries	Antenatal care intermittent presumptive treatment for malaria	Antenatal care	Yes	Health provider quality	Cross-sectional	At least 4 antenatal visits	Quality was measured at facility level
				Antenatal care Syphilis screening	Antenatal care syphilis screening	Yes	Health provider quality	Cross-sectional	At least 4 antenatal visits	Facility categorised as high quality

				Antenatal care tetanus toxoid vaccination	Antenatal care tetanus toxoid vaccination	Yes	Health provider quality	Cross-sectional	At least 4 antenatal visits	Facility categorised as high quality
				Antenatal care hypertensive case management	Antenatal hypertensive case management	Yes	Health provider quality	Cross-sectional	At least 4 antenatal visits	Facility categorised as high quality
				Antenatal care use of iron and folic acid	Antenatal care Iron and Folic Acid use	Yes	Health provider quality	Cross-sectional	At least 4 antenatal visits	Facility categorised as high quality
101	Khan et al. [145]	2000	India	Measles immunisation	Immunisation	Yes	Biomarker of functionality	Cross-sectional	Parent recall/ card	Seroprevalence
102	Kiey et al. [146]	2012	Zambia	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	At least 4 antenatal care visits with skilled Health workers and receipt of 8 interventions (weight measurement, height measurement, blood pressure measurement, urine sample taken for analysis, blood sample taken for analysis, offered VCT, iron supplementation provided, antimalarial drug provided for IPT, birth preparedness plan dis- cussed, treatment provided for intestinal parasites and tetanus toxoid vaccination)/ facility categorised as High quality
				Antenatal care use of iron and folic acid	Antenatal care Iron and Folic Acid use	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	Folate/iron supplementation given or bought
				Antenatal care tetanus toxoid vaccination	Antenatal care tetanus toxoid vaccination	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	Tetanus vaccination received
				Antenatal care HIV counselling and test	Antenatal care HIV counselling and screening	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	Counselling and testing are offered
				Antenatal care intermittent presumptive treatment for malaria	Antenatal care malaria presumptive intermittent treatment	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	IPT of malaria taken
				Antenatal care Anaemia screening	Antenatal care anaemia screening	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	Blood sample given

				Antenatal care Syphilis screening	Antenatal care syphilis screening	Yes	Content care	Cross-sectional	At least 1 antenatal visit by a skilled health worker	Blood sample given
103	Leyvraz et al. (b) [147]	2016	India	Fortified complementary food	Nutrition prevention	Yes	Content care	Cross-sectional	Child consumes the supplement sometimes or always	Child consumes the supplement always
104	Marchant et al. [148]	2015	India, Nigeria, Ethiopia	Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit	Receipt of 8 interventions (weight and height; blood pressure; urine and blood tests; counselling for breastfeeding, danger signs, birth preparedness)
				Skilled birth attendance	Skilled birth attendance	Yes	Content care	Cross-sectional	Birth attended by skilled health workers	Active management of third stage of labour
				Postnatal care	Postnatal care	Yes	Content care	Cross-sectional	Postpartum check with 48 hours	All 5: breasts and bleeding; counselling on danger signs, nutrition, family planning/ All 5: Weigh new-born, check cord, counsel on breastfeeding, thermal care, danger signs
105	Millar et al. [149]	2014	Nigeria	Malaria treatment	Malaria treatment	Yes	Content care	Cross-sectional	Treatment was sought outside of the home	Diagnostic blood test/ received a prompt ACT
106	Mokdad et al. [150]	2015	El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama	Measles, mumps, rubella immunisation	Immunisation	Yes	Content care	Cross-sectional	Recall and card	Recall and card discounting missed opportunities
107	Nanthavong et al. [151]	2015	Lao	Tetanus immunisation	Immunisation	Yes	Biomarker of funcionality	Cross-sectional	Recall and card	Seroprevalence
				Diphtheria immunisation	Immunisation	Yes	Biomarker of funcionality	Cross-sectional	Recall and card	Seroprevalence
108	Ndyomugyenyi et al. [152]	2010	Uganda	Antenatal care intermittent presumptive treatment for malaria	Antenatal care malaria presumptive intermittent treatment	Yes	Content care	Cross-sectional	At least 1 antenatal visit	IPT of malaria received
109	Smith et al. [153]	2010	Senegal	Malaria treatment	Malaria treatment	Yes	Content care	Cross-sectional	Treatment was sought	Received artemisinin-combination therapy (ACT) within first 48 hours
110	Wangdi et al. [154]	2014	Bhutan	Malaria vector control: LLIN (long-lasting	Malaria vector control (ITN)	Yes	Content care	Cross-sectional	Owner of a LLITN	Use a LLITN

				insecticide-treated nets) intervention						
111	Bekuma et al. [155]	2019	Ethiopia	Mass drug administration for Trachoma	Trachoma MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug (swallow)
112	Eboreime et al. [156]	2019	Nigeria	Immunisation Programme	Immunisation	Yes	Content care	Cross-sectional	Received first pentavalent vaccine in the previous year (Children <12m)	Complete the whole programme in the previous year (Children <12m)
				Malaria treatment (children)	Malaria treatment	Yes	Content care	Cross-sectional	Having fever and using health facility services (Children <60m)	Rapid diagnostic tests and prompt treatment with ACT (Children <60m) with ACT (Children <60m)
				Antenatal care	Antenatal care	Yes	Content care	Cross-sectional	At least 1 antenatal visit	Attended ANC clinic during the pregnancy four times
				Skilled birth attendance	Skilled birth attendance	Yes	Content care	Cross-sectional	Deliveries in health facilities	Received postnatal check-up within 48hours at health facilities
113	Hashiguchi et al. [157]	2019	Japan	Treatment of diabetes	Diabetes treatment	Yes	Biomarker of funcionality	Cross-sectional	Diabetes according to registers	Reduction in fasting plasma glucose compared with treatment targets
				Treatment of hypertension	Hypertension treatment	Yes	Biomarker of funcionality	Cross-sectional	Hypertension according to registers	Reduction in systolic blood pressure compared with treatment targets
				Treatment of hyperlipidaemia	Dyslipidaemia treatment	Yes	Biomarker of funcionality	Cross-sectional	Hyperlipidaemia according to registers	Reduction in total cholesterol compared with treatment targets
114	Jordans et al. [51]	2019	Nepal	Treatment of depression	Mental health: Depression	Yes	Functionality	Cross-sectional/ Cohort	-	Minimally adequate treatment/ PHQ-9/ WHODAS
				Treatment of alcohol use disorder	Mental health: Drug and substance use disorder	Yes	Functionality	Cross-sectional/ Cohort	-	Minimally adequate treatment/ SIP-2R/ AUDIT
				Treatment of psychosis	Mental health: Psychosis and severe mental disorders	No	Functionality	Cross-sectional/ Cohort	Extracted from records of population admitted	PANSS/ WHODAS
				Treatment of epilepsy	Mental health: Epilepsy	No	Functionality	Cross-sectional/ Cohort	Extracted from records of population admitted	Number of seizures/ WHODAS
115	Leslie et al. [158]	2019	México	Antenatal care	Antenatal care	No	Incidence of disease	Cross-sectional	At least 1 antenatal visit	Full term delivery

				Skilled birth attendance	Skilled birth attendance	No	Incidence of disease	Cross-sectional	Deliveries in healthcare facilities	Delivery without complication or death
				Postnatal care	Postnatal care	No	Incidence of disease	Cross-sectional	Deliveries (live new-born) in healthcare facilities	live new-born alive at 28th day with-out death due to respiratory infection, nosocomial infection or sepsis
				Diarrhoea (child/severe)	Diarrhoea	No	Incidence of disease	Cross-sectional	Attendance to healthcare provider	Did not result in hospitalization
				Acute respiratory infections (child)	Acute respiratory infections	No	Incidence of disease	Cross-sectional	Health care contact	Did not result in hospitalization
				Treatment of hypertension	Hypertension treatment	No	Biomarker of funcionality	Cross-sectional	Attendance to hypertension measurement last year/ no hospitalization related to hypertension	Reduction in systolic blood pressure compared with treatment targets
				Treatment of diabetes	Diabetes treatment	No	Biomarker of funcionality	Cross-sectional	Attendance to healthcare provider in previous last year/ no hospitalization related to diabetes	Reduction in fasting plasma glucose compared with treatment targets
				Breast cancer treatment	Breast cancer screening	No	Content care	Cross-sectional	Diagnosed cases of breast cancer	Confirmation in 30 days and treatment before 21 days after diagnosis
				Cervical cancer treatment	Cervical cancer screening	No	Content care	Cross-sectional	Diagnosed cases of cervical cancer	Confirmation in 30 days and treatment before 21 days after diagnosis
116	Chang et al. [159]	2019	Haiti	Malaria treatment	Malaria treatment	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
117	Mulebeke et al. [160]	2019	Uganda	Mass drug administration for malaria	Malaria vector control (MDA)	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
118	Ntambi at al. [161]	2019	Somalia	Acute malnutrition management (child)	Nutrition treatment	Yes	Functionality	Cross-sectional	Admission to malnutrition service	Cured from malnutrition service
119	Kulkarni et al. [162]	2019	India	Mass drug administration for Lymphatic Filariasis	Lymphatic Filariasis MDA	Yes	Content care	Cross-sectional	Received the drug	Consume the drug
120	Servan-Mori et al. [163]	2019	México	Maternal and child health service	Maternal and child health service	Yes	Content care	Cohort	Skilled ANC + initial visit during the first trimester + at least 4 visits + adequate content of ANC (8 procedures) + delivery in healthcare facility + skilled delivery	No maternal complication in childbirth + normal child weight
121	Smith et al. [164]	2019	-	Malaria treatment	Malaria treatment	Yes	Content care	Simulation methods	-	Treatment efficacy = 100% + given within 14 days

122	Tsiachristas et al. [165]	2019	Kenya	Newburn care	Postnatal care	Yes	Content care	Cost-effectiveness analysis	-	Performed at least 80% of the required tasks per care category according to references
123	Venkateswaran et al. [166]	2019	Palestine	Antenatal care Blood pressure measurement	Antenatal care blood pressure measurement	Yes	Content care	Cross-sectional	At least one blood pressure measurement	in all recommended ANC visits
				Antenatal care Symphysis-fundal height	Antenatal care measurement of symphysis-fundal height	Yes	Content care	Cross-sectional	At least one SFH measurement	in all recommended ANC visits
				Antenatal care Anaemia screening	Antenatal care anaemia screening	Yes	Content care	Cross-sectional	At least one haemoglobin determination	at first visit, 24-28 and 36 weeks
				Antenatal care Ultrasound	Antenatal care Ultrasound	Yes	Content care	Cross-sectional	At least one ultrasound examination	at first visit, 24-28 and 36 weeks
				Antenatal care Gestational Diabetes screening	Antenatal care gestational diabetes screening	Yes	Content care	Cross-sectional	At least one determination	Urine sugar test at booking and blood sugar test at 24–28 weeks
				Antenatal care Abacteriuria Asymptomatic screening	Antenatal care asymptomatic bacteriuria	Yes	Content care	Cross-sectional	Urine microscopy test	Urine microscopy test in the first visit
				Antenatal care Rh type screening	Antenatal care Rh type screening	Yes	Content care	Cross-sectional	Rh-typing	Rh-typing in the first visit
				Antenatal care tetanus toxoid vaccination	Antenatal care tetanus toxoid vaccination	Yes	Content care	Cross-sectional	Status is checked by asking for history of immunization or reviewing records	Status is checked by asking for history of immunization or reviewing records in the first visit
124	Wang et al. [167]	2019	Bangladesh, Haiti, Malawi, Nepal, Senegal, Tanzania	Skilled birth attendance	Skilled birth attendance	Yes	Health provider quality	Cross-sectional	Birth attended in a health facility	Facility categorised as high quality
125	Yakob et al. [168]	2019	Ethiopia	Family planning	Family planning	Yes	Health provider quality	Cross-sectional	Use of a modern contraceptive method	Use a facility index of quality
				Antenatal care	Antenatal care	Yes	Health provider quality	Cross-sectional	At least 1 antenatal visit by skilled professional	Use a facility index of quality
126	Fuseini et al. [169]	2020	Equatorial Guinea	Malaria vector control: IRS (indoor residual spraying) intervention	Malaria vector control (IRS)	Yes	Content care	Cross-sectional	-	The concertation of insecticide was measured in the households
127	Mahdavi et al. [170]	2020	Iran	Treatment of hypertension	Hypertension treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets

128	Sarma et al. [171]	2020	Bangladesh	Fortified complementary foods (micronutrient powders) for children	Nutrition prevention	Yes	Content care	Cross-sectional	Child ever been fed the product	Child ever been fed the product in the previous 7 days at least in 3 days
129	Zhao et al. [172]	2020	China	Treatment of hypertension	Hypertension treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets
130	Aguilera et al. [52]	2014	Chile	Treatment of hypertension	Hypertension treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets
				Treatment of diabetes	Diabetes treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of oral hypoglycaemics or insulin	Reduction in fasting plasma glucose compared with treatment targets
131	Huda et al. [53]	2014	Bangladesh	Treatment of hypertension	Hypertension treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of antihypertensive agents	Reduction in systolic blood pressure compared with treatment targets
				Treatment of diabetes	Diabetes treatment	Yes	Biomarker of functionality	Cross-sectional	Self-reported use of oral hypoglycaemics or insulin	Reduction in fasting plasma glucose compared with treatment targets

Supplementary material chapter 2

Contents

Table S2.1. Health State Description - World Health Survey questionnaire

Overall Health

The first questions are about your overall health, including both your physical and your mental health.

1	In general, how would you rate your health today?	Very good	Good	Moderate	Bad	Very Bad
2	Overall in the last 30 days, how much difficulty did you have with work or household activities?	None	Mild	Moderate	Severe	Extreme/ Cannot do

Now I would like to review different functions of your body. When answering these questions, I would like you to think about the last 30 days, taking both good and bad days into account. When I ask about difficulty, I would like you to consider how much difficulty you have had, on an average, in the past 30 days, while doing the activity in the way that you usually do it. By difficulty I mean requiring increased effort, discomfort or pain, slowness or changes in the way you do the activity. Please answer this question taking into account any assistance you have available. (Read and show scale to respondent).

Mobility

3	Overall in the last 30 days, how much difficulty did you have with moving around?	None	Mild	Moderate	Severe	Extreme/ Cannot do
4	In the last 30 days, how much difficulty did you have in vigorous activities, such as running 3 km (or equivalent) or cycling?	None	Mild	Moderate	Severe	Extreme/ Cannot do

Self Care

5	Overall in the last 30 days, how much difficulty did you have with self- care, such as washing or dressing yourself?	None	Mild	Moderate	Severe	Extreme/ Cannot do
6	In the last 30 days, how much difficulty did you have in taking care of and maintaining your general appearance (e.g. grooming, looking neat and tidy etc.)	None	Mild	Moderate	Severe	Extreme/ Cannot do

Pain and Discomfort

7	Overall in the last 30 days, how much of bodily aches or pains did you have?	None	Mild	Moderate	Severe	Extreme
8	In the last 30 days, how much bodily discomfort did you have?	None	Mild	Moderate	Severe	Extreme

Cognition

9	Overall in the last 30 days, how much difficulty did you have with concentrating or remembering things?	None	Mild	Moderate	Severe	Extreme/ Cannot do
10	In the last 30 days, how much difficulty did you have in learning a new task (for example, learning how to get to a new place, learning a new game, learning a new recipe etc.)?	None	Mild	Moderate	Severe	Extreme/ Cannot do

Interpersonal activities

11	Overall in the last 30 days, how much difficulty did you have with personal relationship or participation in the community?	None	Mild	Moderate	Severe	Extreme/ Cannot do
12	In the last 30 days, how much difficulty did you have in dealing with conflicts and tensions with others?	None	Mild	Moderate	Severe	Extreme/ Cannot do

Vision

	Do you wear glasses or contact lenses? (If Respondent says YES to this question, preface the next 2 questions with "Please answer the following questions taking into account your glasses or contact lenses".)	Yes			No	
13	In the last 30 days, how much difficulty did you have in seeing and recognizing a person you know across the road (i.e. from a distance of about 20 meters)?	None	Mild	Moderate	Severe	Extreme/ Cannot do
14	In the last 30 days, how much difficulty did you have in seeing and recognizing an object at arm's length or in reading?	None	Mild	Moderate	Severe	Extreme/ Cannot do

Sleep and Energy

15	Overall in the last 30 days, how much of a problem did you have with sleeping, such as falling asleep, waking up frequently during the night or waking up too early in the morning?	None	Mild	Moderate	Severe	Extreme
16	In the last 30 days, how much of a problem did you have due to not feeling rested and refreshed during the day (e.g. feeling tired, not having energy)?	None	Mild	Moderate	Severe	Extreme

Affect

17	Overall in the last 30 days, how much of a problem did you have with feeling sad, low or depressed?	None	Mild	Moderate	Severe	Extreme
18	Overall in the last 30 days, how much of a problem did you have with worry or anxiety?	None	Mild	Moderate	Severe	Extreme

Table S2.2. Polychoric correlation matrix between items of the HSQ-WHS questionnaire of the sample (n=2,646). Chilean National Health Survey 2009-2010.

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18
I1	1,00																	
I2	0,64	1,00																
I3	0,59	0,77	1,00															
I4	0,43	0,54	0,63	1,00														
I5	0,54	0,70	0,74	0,53	1,00													
I6	0,50	0,67	0,70	0,47	0,92	1,00												
I7	0,53	0,59	0,57	0,46	0,48	0,49	1,00											
I8	0,56	0,63	0,60	0,45	0,52	0,47	0,84	1,00										
I9	0,41	0,46	0,42	0,33	0,43	0,47	0,37	0,39	1,00									
I10	0,43	0,50	0,50	0,40	0,54	0,55	0,36	0,38	0,73	1,00								
I11	0,46	0,53	0,53	0,38	0,57	0,62	0,38	0,41	0,50	0,58	1,00							
I12	0,39	0,44	0,42	0,28	0,45	0,50	0,33	0,35	0,45	0,47	0,64	1,00						
I13	0,34	0,39	0,41	0,33	0,42	0,44	0,29	0,30	0,35	0,38	0,35	0,26	1,00					
I14	0,35	0,37	0,40	0,34	0,43	0,42	0,30	0,32	0,31	0,37	0,31	0,24	0,68	1,00				
I15	0,39	0,45	0,43	0,33	0,38	0,40	0,43	0,45	0,40	0,37	0,39	0,36	0,31	0,31	1,00			
I16	0,41	0,44	0,38	0,27	0,36	0,38	0,41	0,44	0,41	0,35	0,40	0,42	0,30	0,26	0,65	1,00		
I17	0,52	0,52	0,43	0,33	0,44	0,45	0,47	0,47	0,45	0,39	0,45	0,45	0,30	0,30	0,53	0,54	1,00	
I18	0,43	0,44	0,35	0,26	0,32	0,34	0,45	0,44	0,44	0,31	0,38	0,41	0,26	0,24	0,51	0,55	0,72	1,00

Figure S2.1. Plot of the polychoric correlation matrix between items of the HSQ-WHS questionnaire of the sample (n=2,646). Chilean National Health Survey 2009-2010.

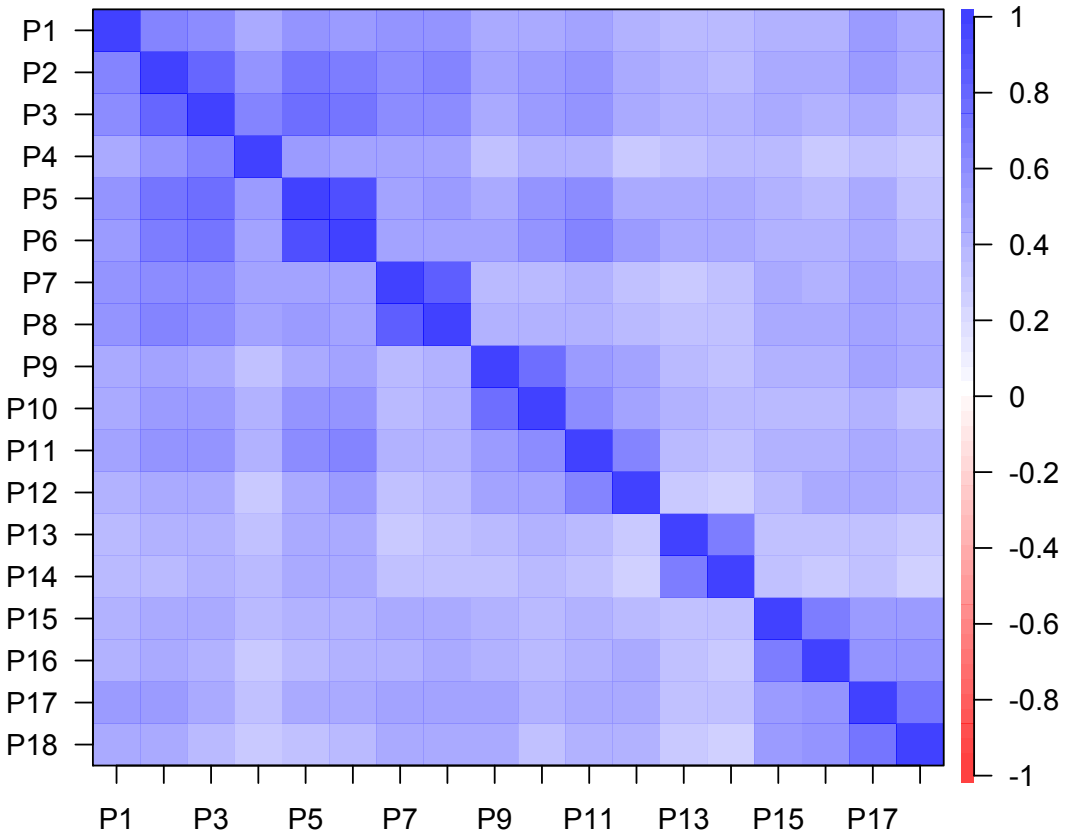


Table S2.3. Coefficients of Confirmatory Factor Analysis for the Health State Description questionnaire of the World Health Survey (HSD-WHS). Chilean National Health Survey 2009/2010 (n=5,293)

Note:
Syntaxis for R using lavaan

```
Equ=        'GHealth=~ P1 + P2
            Movi=~ P3 + P4
            SelfC=~ P5 + P6
            Pain=~ P7 + P8
            Cogni=~ P9 +P10
            Interp=~ P11 + P12
            Vision=~ P13 + P14
            Energ=~ P15 + P16
            Affect=~ P17 + P18
            Disab=~ GHealth + Movi + SelfC + Pain + Cogni + Interp + Vision + Energ + Affect'

Model<- lavaan::cfa(Equ, data, orthogonal=TRUE, ordered=paste0("P", 1:18),
parametrization="theta")
```

	lhs	op	rhs	est	se	z	pvalue	ci.lower	ci.upper
1	GHealth	=~	P1	1.00	0.00	NA	NA	1.00	1.00
2	GHealth	=~	P2	1.18	0.02	52.84	0.00	1.13	1.22
3	Movi	=~	P3	1.00	0.00	NA	NA	1.00	1.00
4	Movi	=~	P4	0.73	0.02	36.33	0.00	0.69	0.77
5	SelfC	=~	P5	1.00	0.00	NA	NA	1.00	1.00
6	SelfC	=~	P6	0.97	0.02	49.26	0.00	0.93	1.01
7	Pain	=~	P7	1.00	0.00	NA	NA	1.00	1.00
8	Pain	=~	P8	1.04	0.02	68.92	0.00	1.01	1.07
9	Cogni	=~	P9	1.00	0.00	NA	NA	1.00	1.00
10	Cogni	=~	P10	1.03	0.02	44.31	0.00	0.99	1.08
11	Interp	=~	P11	1.00	0.00	NA	NA	1.00	1.00
12	Interp	=~	P12	0.88	0.03	32.88	0.00	0.83	0.93
13	Vision	=~	P13	1.00	0.00	NA	NA	1.00	1.00
14	Vision	=~	P14	0.94	0.04	25.56	0.00	0.87	1.01
15	Energ	=~	P15	1.00	0.00	NA	NA	1.00	1.00
16	Energ	=~	P16	0.95	0.02	38.18	0.00	0.90	1.00
17	Affect	=~	P17	1.00	0.00	NA	NA	1.00	1.00
18	Affect	=~	P18	0.91	0.02	52.49	0.00	0.88	0.94
19	Disab	=~	GHealth	1.00	0.00	NA	NA	1.00	1.00
20	Disab	=~	Movi	1.13	0.02	47.46	0.00	1.08	1.18
21	Disab	=~	SelfC	1.08	0.03	34.68	0.00	1.02	1.14
22	Disab	=~	Pain	0.94	0.02	42.82	0.00	0.90	0.98
23	Disab	=~	Cogni	0.86	0.02	35.99	0.00	0.82	0.91
24	Disab	=~	Interp	0.95	0.03	34.16	0.00	0.89	1.00
25	Disab	=~	Vision	0.70	0.03	24.30	0.00	0.65	0.76
26	Disab	=~	Energ	0.89	0.03	35.43	0.00	0.84	0.94
27	Disab	=~	Affect	0.96	0.02	43.13	0.00	0.91	1.00
28	P1	|	t1	-1.31	0.03	-38.87	0.00	-1.38	-1.25
29	P1	|	t2	0.17	0.02	6.90	0.00	0.12	0.22
30	P1	|	t3	1.46	0.04	39.90	0.00	1.39	1.53
31	P1	|	t4	2.39	0.08	30.76	0.00	2.24	2.55
	P2	|	t1	0.08	0.02	3.40	0.00	0.04	0.13
33	P2	|	t2	0.65	0.03	24.82	0.00	0.60	0.71
34	P2	|	t3	1.41	0.04	39.63	0.00	1.34	1.48
35	P2	|	t4	2.12	0.06	35.57	0.00	2.00	2.24
36	P3	|	t1	0.52	0.03	20.36	0.00	0.47	0.57
37	P3	|	t2	0.90	0.03	31.90	0.00	0.85	0.96
38	P3	|	t3	1.47	0.04	39.93	0.00	1.40	1.54

39	P3	|	t4	2.16	0.06	34.98	0.00	2.04	2.28
40	P4	|	t1	-0.20	0.02	-8.21	0.00	-0.25	-0.15
41	P4	|	t2	0.11	0.02	4.49	0.00	0.06	0.16
42	P4	|	t3	0.46	0.03	18.10	0.00	0.41	0.51
43	P4	|	t4	0.90	0.03	31.69	0.00	0.84	0.95
44	P5	|	t1	1.32	0.03	38.93	0.00	1.25	1.38
45	P5	|	t2	1.56	0.04	40.12	0.00	1.48	1.63
46	P5	|	t3	1.90	0.05	38.40	0.00	1.80	2.00
47	P5	|	t4	2.45	0.08	29.72	0.00	2.29	2.61
48	P6	|	t1	1.33	0.03	39.02	0.00	1.26	1.39
49	P6	|	t2	1.62	0.04	40.09	0.00	1.54	1.70
50	P6	|	t3	1.99	0.05	37.32	0.00	1.89	2.10
51	P6	|	t4	2.45	0.08	29.72	0.00	2.29	2.61
52	P7	|	t1	-0.49	0.03	-19.29	0.00	-0.54	-0.44
53	P7	|	t2	0.24	0.02	9.73	0.00	0.19	0.29
54	P7	|	t3	0.99	0.03	33.84	0.00	0.93	1.05
55	P7	|	t4	1.98	0.05	37.49	0.00	1.88	2.08
56	P8	|	t1	-0.49	0.03	-19.29	0.00	-0.54	-0.44
57	P8	|	t2	0.24	0.02	9.92	0.00	0.20	0.29
58	P8	|	t3	1.02	0.03	34.39	0.00	0.96	1.07
59	P8	|	t4	1.95	0.05	37.80	0.00	1.85	2.06
60	P9	|	t1	-0.13	0.02	-5.42	0.00	-0.18	-0.08
61	P9	|	t2	0.55	0.03	21.27	0.00	0.50	0.60
62	P9	|	t3	1.28	0.03	38.58	0.00	1.22	1.35
63	P9	|	t4	2.25	0.07	33.35	0.00	2.12	2.39
64	P10	|	t1	0.32	0.02	12.75	0.00	0.27	0.36
65	P10	|	t2	0.87	0.03	30.99	0.00	0.81	0.92
66	P10	|	t3	1.48	0.04	39.95	0.00	1.40	1.55
67	P10	|	t4	2.28	0.07	32.92	0.00	2.14	2.41
68	P11	|	t1	0.69	0.03	25.86	0.00	0.64	0.74
69	P11	|	t2	1.14	0.03	36.74	0.00	1.08	1.21
70	P11	|	t3	1.62	0.04	40.09	0.00	1.54	1.70
71	P11	|	t4	2.17	0.06	34.83	0.00	2.05	2.29
72	P12	|	t1	0.39	0.03	15.60	0.00	0.34	0.44
73	P12	|	t2	0.93	0.03	32.45	0.00	0.87	0.98
74	P12	|	t3	1.52	0.04	40.06	0.00	1.44	1.59
75	P12	|	t4	2.28	0.07	32.92	0.00	2.14	2.41
76	P13	|	t1	0.37	0.02	14.99	0.00	0.33	0.42
77	P13	|	t2	0.78	0.03	28.66	0.00	0.73	0.84
78	P13	|	t3	1.25	0.03	38.25	0.00	1.19	1.32
79	P13	|	t4	2.11	0.06	35.70	0.00	2.00	2.23
80	P14	|	t1	0.32	0.02	12.75	0.00	0.27	0.36
81	P14	|	t2	0.71	0.03	26.60	0.00	0.66	0.76
82	P14	|	t3	1.19	0.03	37.44	0.00	1.13	1.25
83	P14	|	t4	2.16	0.06	34.98	0.00	2.04	2.28
84	P15	|	t1	-0.10	0.02	-4.10	0.00	-0.15	-0.05
85	P15	|	t2	0.37	0.02	14.95	0.00	0.32	0.42
86	P15	|	t3	0.96	0.03	33.27	0.00	0.91	1.02
87	P15	|	t4	1.89	0.05	38.51	0.00	1.79	1.98
88	P16	|	t1	-0.12	0.02	-4.99	0.00	-0.17	-0.07
89	P16	|	t2	0.55	0.03	21.27	0.00	0.50	0.60
90	P16	|	t3	1.22	0.03	37.82	0.00	1.16	1.28
91	P16	|	t4	2.17	0.06	34.83	0.00	2.05	2.29
92	P17	|	t1	-0.32	0.02	-12.86	0.00	-0.37	-0.27
93	P17	|	t2	0.46	0.03	18.22	0.00	0.41	0.51
94	P17	|	t3	1.14	0.03	36.60	0.00	1.07	1.20
95	P17	|	t4	1.99	0.05	37.41	0.00	1.88	2.09
96	P18	|	t1	-0.50	0.03	-19.71	0.00	-0.55	-0.45
97	P18	|	t2	0.25	0.02	10.31	0.00	0.21	0.30
98	P18	|	t3	0.93	0.03	32.45	0.00	0.87	0.98
99	P18	|	t4	1.95	0.05	37.88	0.00	1.85	2.05
100	P1	~~	P1	0.46	0.00	NA	NA	0.46	0.46
101	P2	~~	P2	0.25	0.00	NA	NA	0.25	0.25
102	P3	~~	P3	0.16	0.00	NA	NA	0.16	0.16
103	P4	~~	P4	0.56	0.00	NA	NA	0.56	0.56
104	P5	~~	P5	0.04	0.00	NA	NA	0.04	0.04
105	P6	~~	P6	0.10	0.00	NA	NA	0.10	0.10

106	P7	~~	P7	0.19	0.00	NA	NA	0.19	0.19
107	P8	~~	P8	0.13	0.00	NA	NA	0.13	0.13
108	P9	~~	P9	0.29	0.00	NA	NA	0.29	0.29
109	P10	~~	P10	0.24	0.00	NA	NA	0.24	0.24
110	P11	~~	P11	0.29	0.00	NA	NA	0.29	0.29
111	P12	~~	P12	0.44	0.00	NA	NA	0.44	0.44
112	P13	~~	P13	0.28	0.00	NA	NA	0.28	0.28
113	P14	~~	P14	0.36	0.00	NA	NA	0.36	0.36
114	P15	~~	P15	0.30	0.00	NA	NA	0.30	0.30
115	P16	~~	P16	0.36	0.00	NA	NA	0.36	0.36
116	P17	~~	P17	0.21	0.00	NA	NA	0.21	0.21
117	P18	~~	P18	0.34	0.00	NA	NA	0.34	0.34
118	GHealth	~~	GHealth	0.03	0.01	3.08	0.00	0.01	0.05
119	Movi	~~	Movi	0.18	0.02	7.97	0.00	0.14	0.22
120	SelfC	~~	SelfC	0.36	0.03	12.84	0.00	0.30	0.41
121	Pain	~~	Pain	0.35	0.01	25.99	0.00	0.32	0.38
122	Cogni	~~	Cogni	0.32	0.02	19.30	0.00	0.29	0.36
123	Interp	~~	Interp	0.25	0.02	12.40	0.00	0.21	0.29
124	Vision	~~	Vision	0.47	0.03	17.56	0.00	0.42	0.52
125	Energ	~~	Energ	0.29	0.02	16.56	0.00	0.26	0.33
126	Affect	~~	Affect	0.32	0.02	18.83	0.00	0.29	0.36
127	Disab	~~	Disab	0.51	0.02	28.59	0.00	0.48	0.55
128	P1	~*~	P1	1.00	0.00	NA	NA	1.00	1.00
129	P2	~*~	P2	1.00	0.00	NA	NA	1.00	1.00
130	P3	~*~	P3	1.00	0.00	NA	NA	1.00	1.00
131	P4	~*~	P4	1.00	0.00	NA	NA	1.00	1.00
132	P5	~*~	P5	1.00	0.00	NA	NA	1.00	1.00
133	P6	~*~	P6	1.00	0.00	NA	NA	1.00	1.00
134	P7	~*~	P7	1.00	0.00	NA	NA	1.00	1.00
135	P8	~*~	P8	1.00	0.00	NA	NA	1.00	1.00
136	P9	~*~	P9	1.00	0.00	NA	NA	1.00	1.00
137	P10	~*~	P10	1.00	0.00	NA	NA	1.00	1.00
138	P11	~*~	P11	1.00	0.00	NA	NA	1.00	1.00
139	P12	~*~	P12	1.00	0.00	NA	NA	1.00	1.00
140	P13	~*~	P13	1.00	0.00	NA	NA	1.00	1.00
141	P14	~*~	P14	1.00	0.00	NA	NA	1.00	1.00
142	P15	~*~	P15	1.00	0.00	NA	NA	1.00	1.00
143	P16	~*~	P16	1.00	0.00	NA	NA	1.00	1.00
144	P17	~*~	P17	1.00	0.00	NA	NA	1.00	1.00
145	P18	~*~	P18	1.00	0.00	NA	NA	1.00	1.00
146	P1	~1		0.00	0.00	NA	NA	0.00	0.00
147	P2	~1		0.00	0.00	NA	NA	0.00	0.00
148	P3	~1		0.00	0.00	NA	NA	0.00	0.00
149	P4	~1		0.00	0.00	NA	NA	0.00	0.00
150	P5	~1		0.00	0.00	NA	NA	0.00	0.00
151	P6	~1		0.00	0.00	NA	NA	0.00	0.00
152	P7	~1		0.00	0.00	NA	NA	0.00	0.00
153	P8	~1		0.00	0.00	NA	NA	0.00	0.00
154	P9	~1		0.00	0.00	NA	NA	0.00	0.00
155	P10	~1		0.00	0.00	NA	NA	0.00	0.00
156	P11	~1		0.00	0.00	NA	NA	0.00	0.00
157	P12	~1		0.00	0.00	NA	NA	0.00	0.00
158	P13	~1		0.00	0.00	NA	NA	0.00	0.00
159	P14	~1		0.00	0.00	NA	NA	0.00	0.00
160	P15	~1		0.00	0.00	NA	NA	0.00	0.00
161	P16	~1		0.00	0.00	NA	NA	0.00	0.00
162	P17	~1		0.00	0.00	NA	NA	0.00	0.00
163	P18	~1		0.00	0.00	NA	NA	0.00	0.00
164	GHealth	~1		0.00	0.00	NA	NA	0.00	0.00
165	Movi	~1		0.00	0.00	NA	NA	0.00	0.00
166	SelfC	~1		0.00	0.00	NA	NA	0.00	0.00
167	Pain	~1		0.00	0.00	NA	NA	0.00	0.00
168	Cogni	~1		0.00	0.00	NA	NA	0.00	0.00
169	Interp	~1		0.00	0.00	NA	NA	0.00	0.00
170	Vision	~1		0.00	0.00	NA	NA	0.00	0.00
171	Energ	~1		0.00	0.00	NA	NA	0.00	0.00
172	Affect	~1		0.00	0.00	NA	NA	0.00	0.00

173	Disab	~1		0.00	0.00	NA	NA	0.00	0.00

lhs: left hand side of the operator/ rhs: right hand side of the operator/ op: operator/ est: estimate/ se: standard error

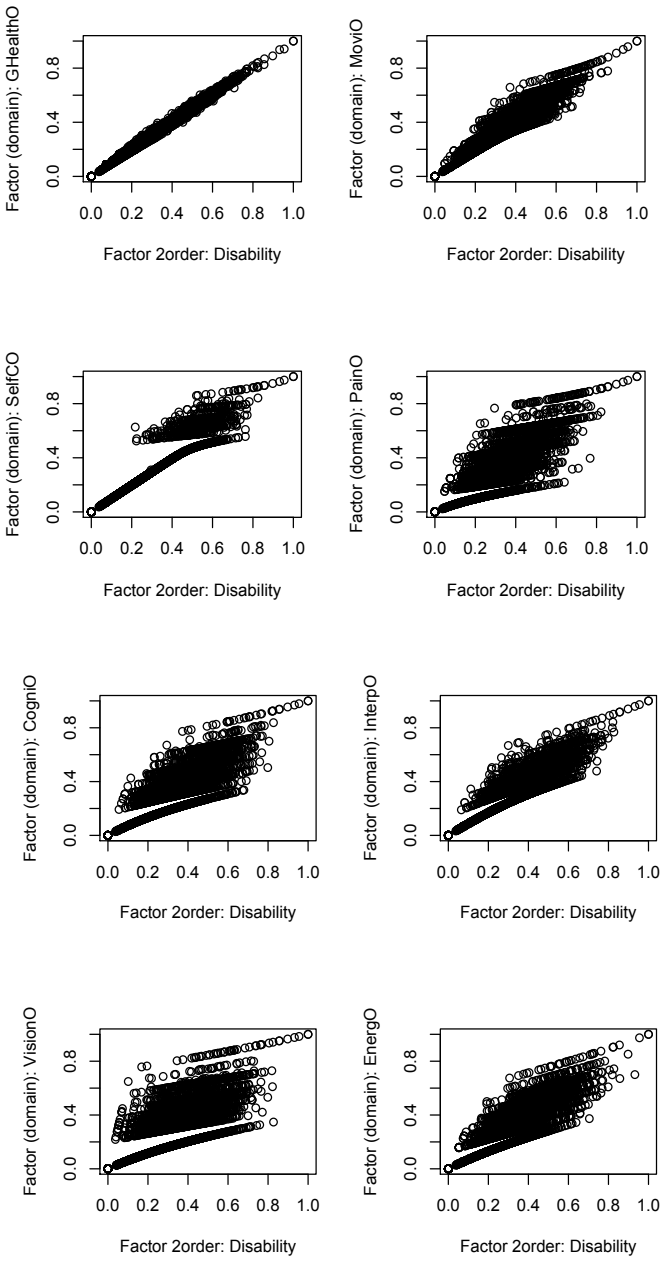
Disab: disability/ Movi: mobility domain/ SelfC: self-care domain/ Pain: pain and discomfort domain/ Cgni: cognition domain/ Interp: interpersonal activities domain/ Vision: vision domain/ Energ: sleep and energy domain/ Affect: affect domain/ GHealth: overall health domain/ P1-P18: represent the 18 items of the questionnaire.

Table S2.4. Correlation matrix between factor scores from the HSQ-WHS questionnaire (n=5,293), using Confirmatory Factor Analysis, respecting categorical nature of data. Chilean National Health Survey 2009-2010.

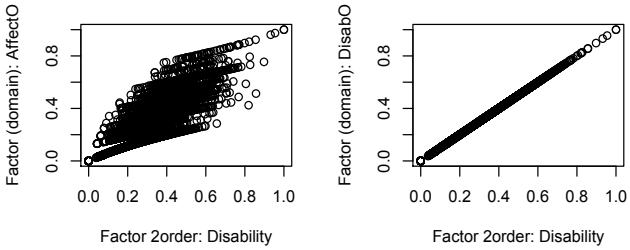
	GHealth	Movi	SelfC	Pain	Cogni	Interp	Vision	Energ	Affect	Disab
GHealth	1.00									
Movi	0.95	1.00								
SelfC	0.95	0.91	1.00							
Pain	0.82	0.78	0.77	1.00						
Cogni	0.81	0.76	0.78	0.62	1.00					
Interp	0.89	0.83	0.86	0.69	0.77	1.00				
Vision	0.69	0.66	0.66	0.53	0.59	0.61	1.00			
Energ	0.84	0.78	0.80	0.69	0.69	0.76	0.58	1.00		
Affect	0.82	0.74	0.77	0.66	0.67	0.74	0.54	0.77	1.00	
Disab	1.00	0.95	0.95	0.83	0.83	0.90	0.70	0.86	0.83	1.00

Disab: disability/ Movi: mobility domain/ SelfC: self-care domain/ Pain: pain and discomfort domain/ Cgni: cognition domain/ Interp: interpersonal activities domain/ Vision: vision domain/ Energ: sleep and energy domain/ Affect: affect domain/ GHealth: overall health domain.

Figure S2.2. Scatterplot of correlation between factor scores from different domains and second order factor score of disability. HSQ-WHS questionnaire (n=5,293), using Confirmatory Factor Analysis, respecting categorical nature of data. Chilean National Health Survey 2009-2010.



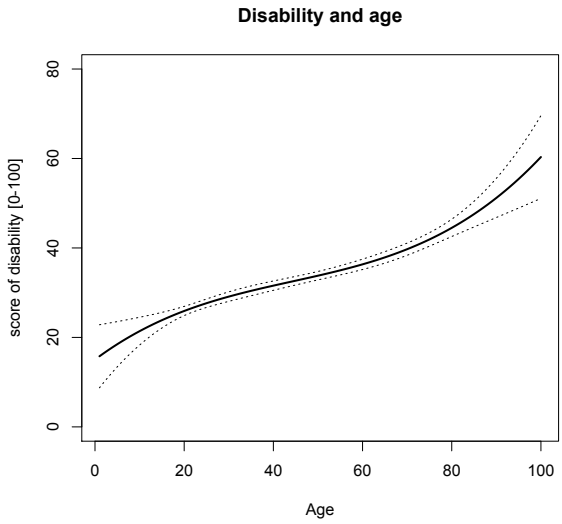
DisabO: disability/ MoviO: mobility domain/ SelfCO: self-care domain/ PainO: pain and discomfort domain/ CgniO: cognition domain/ InterpO: interpersonal activities domain/ VisionO: vision domain/ EnergO: sleep and energy domain/ AffectO: affect domain/ GHealthO: overall health domain.



DisabO: disability/ MoviO: mobility domain/ SelfCO: self-care domain/ PainO: pain and discomfort domain/ CgniO: cognition domain/ InterpO: interpersonal activities domain/ VisionO: vision domain/ EnergO: sleep and energy domain/ AffectO: affect domain/ GHealthO: overall health domain.

Figure S2.3. Predicted values of Disability (panel A) and Domains of Disability (panel B) according to Age. Chilean National Health Survey 2009-2010. (n=5,293), using Confirmatory Factor Analysis, respecting categorical nature of data.

Panel A



Panel B

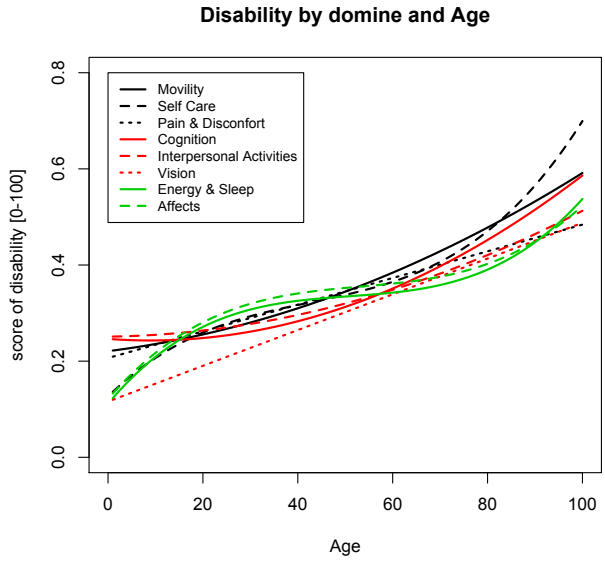


Table S2.5. Results of multivariate regression models on Disability and Domains of Disability. Chilean National Health Survey 2009-2010. (n=4,600)
Disability score and its Domains were calculated using a Confirmatory Factor Analysis, respecting categorical nature of data.

	Disability			Overall Health			Mobility		
	Coeff	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**
Intercept	21,5	18,8	24,2	21,8	19,1	24,5	20,4	17,4	23,4
Sex (females)	4,6	3,4	5,9	4,7	3,4	5,9	4,7	3,4	6,0
Age (10 years)	1,9	1,5	2,3	1,9	1,5	2,3	2,3	1,9	2,7
Education (< 8 years)	0,0	-	-	0,0	-	-	0,0	-	-
Education (8-12 years)	-3,1	-4,9	-1,4	-3,2	-4,9	-1,5	-3,5	-5,5	-1,6
Education (>12 years)	-5,2	-7,4	-3,1	-5,6	-7,7	-3,4	-5,8	-8,1	-3,5
Depression	13,6	12,1	15,2	13,4	11,9	15,0	12,0	10,4	13,7
Hypertension	1,6	0,0	3,3	1,7	0,0	3,3	2,7	0,9	4,4
Diabetes	5,0	2,5	7,4	5,1	2,6	7,5	5,9	3,1	8,8

	Disability			Overall Health			Mobility		
	Coeff	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**
Intercept	21,5	18,8	24,2	21,8	19,1	24,5	20,4	17,4	23,4
Sex (females)	4,6	3,4	5,9	4,7	3,4	5,9	4,7	3,4	6,0
Age (10 years)	1,9	1,5	2,3	1,9	1,5	2,3	2,3	1,9	2,7
Education (< 8 years)	0,0	-	-	0,0	-	-	0,0	-	-
Education (8-12 years)	-3,1	-4,9	-1,4	-3,2	-4,9	-1,5	-3,5	-5,5	-1,6
Education (>12 years)	-5,2	-7,4	-3,1	-5,6	-7,7	-3,4	-5,8	-8,1	-3,5
Depression	13,6	12,1	15,2	13,4	11,9	15,0	12,0	10,4	13,7
Hypertension	1,6	0,0	3,3	1,7	0,0	3,3	2,7	0,9	4,4
Diabetes	5,0	2,5	7,4	5,1	2,6	7,5	5,9	3,1	8,8

	Interpersonal activities			Vision			Spleep & Energy		
	Coeff	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**
Intercept	22,9	19,8	26,0	13,2	10,1	16,3	21,4	18,1	24,7
Sex (females)	2,6	1,1	4,0	5,8	4,3	7,3	4,9	3,2	6,5
Age (10 years)	1,6	1,1	2,1	2,9	2,4	3,3	1,4	0,9	2,0
Education (< 8 years)	0,0	-	-	0,0	-	-	0,0	-	-
Education (8-12 years)	-3,3	-5,2	-1,4	-3,5	-5,6	-1,3	-1,1	-3,1	0,9
Education (>12 years)	-5,5	-7,8	-3,2	-6,2	-8,6	-3,7	-1,0	-3,6	1,5
Depression	14,1	12,3	15,9	10,1	8,2	12,1	15,9	13,9	17,9
Hypertension	1,3	-0,5	3,1	1,8	-0,2	3,8	0,4	-1,7	2,4
Diabetes	4,8	2,1	7,5	4,0	0,5	7,5	3,8	1,0	6,5

	Affect		
	Coeff	**LCI**	**UCI**
Intercept	22,2	18,9	25,5
Sex (females)	5,0	3,4	6,6
Age (10 years)	1,4	0,9	1,9
Education (< 8 years)	0,0	-	-
Education (8-12 years)	-1,6	-3,7	0,6
Education (>12 years)	-1,9	-4,5	0,7
Depression	21,1	19,0	23,1
Hypertension	0,7	-1,3	2,8
Diabetes	4,2	1,3	7,1

Coeff: regression coefficient/ LCI: lower confidence interval/ UCI: upper confidence interval

Table S2.6. Goodness of fit of a confirmatory factor analysis for the Health State Description questionnaire of the World Health Survey (HSD-WHS), assuming continuous variables but including sample weights. Chilean National Health Survey 2009/2010. (n=2,647).

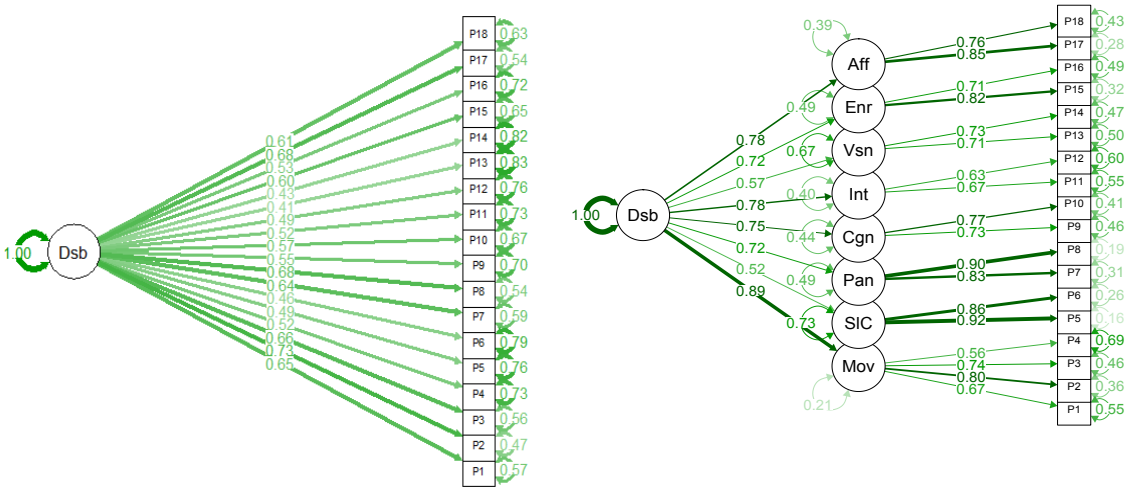
	Unidimentional model	Model suggested by EFA	Theoretical model
χ^2	7127.2	1800.1	1907.2
Df	135	127	126
p value	<0.001	<0.001	<0.001
relative χ^2	52.8	14.2	15.1
TLI	0.61	0.90	0.89
CFI	0.66	0.92	0.91
GFI	0.82	0.95	0.94
RMSEA	0.140	0.071	0.073
RMSEA LCI	0.137	0.068	0.070
RMSA UCI	0.143	0.073	0.076
SRMR	0.085	0.053	0.054

EFA: Exploratory factor analysis / χ^2: chi-square / Df: degree of freedom / TLI: Tucker-Lewis index / CFI: comparative fir index / GFI: goodness of fit index / RMSEA: root mean square error of approximation / SRMR: standardized root mean square residual / LCI: lower 90% confidence interval / UCI: upper 90% confidence interval

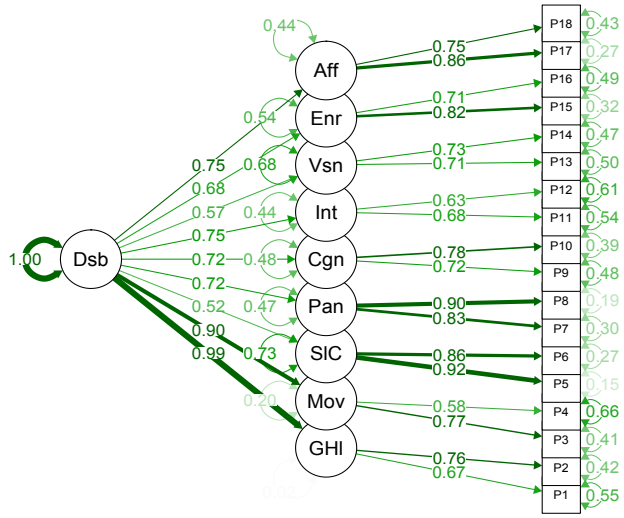
Figure S2.4. Path-diagram of the Health State Description of the World Health Survey (HSD-WHS), assuming continuous variables but including sample weights. Chilean National Health Survey 2009/2010. (n=2,646)

A. Unidimentional model by EFA

B. Model according to the structure suggested



C. Model according to the structure suggested by the theoretical framework



EFA: exploratory factor analysis
Dsb: disability/ Mov: mobility domain/ SLC: self-care domain/ Pan: pain and discomfort domain/ Cgn: cognition domain/ Int: interpersonal activities domain/ Vsn: vision domain/ Enr: sleep and energy domain/ Aff: affect domain/ GHI: Overall Health/ P1-P18: represent the 18 items of the questionnaire.

Values on the straight arrows are the standardized coefficients/ values on the curved arrows are the unexplained variance of the variables.

Table S2.7. Coefficients of Confirmatory Factor Analysis for the Health State Description questionnaire of the World Health Survey (HSD-WHS), assuming continuous variables but including sample weights. Chilean National Health Survey 2009/2010 (n=5,293)

Note:
Syntaxis for R using lavaan

```
Equ=        'GHealth=~ P1 + P2
            Movi=~ P3 + P4
            SelfC=~ P5 + P6
            Pain=~ P7 + P8
            Cogni=~ P9 +P10
            Interp=~ P11 + P12
            Vision=~ P13 + P14
            Energ=~ P15 + P16
            Affect=~ P17 + P18
            Disab=~ GHealth + Movi + SelfC + Pain + Cogni + Interp + Vision + Energ + Affect'

SurvDesign <- svydesign(~1, weights=~sample_weight, data=data)
model<- lavaan::cfa(Equ, data, orthogonal=TRUE)
model<-lavaan.survey(model, SurvDesign)
```

	lhs	op	rhs	est	se	z	pvalue	ci.lower	ci.upper
1	GHealth	=~	P1	1.00	0.00	NA	NA	1.00	1.00
2	GHealth	=~	P2	1.42	0.08	18.19	0.00	1.27	1.58
3	Movi	=~	P3	1.00	0.00	NA	NA	1.00	1.00
4	Movi	=~	P4	1.23	0.07	16.95	0.00	1.08	1.37
5	SelfC	=~	P5	1.00	0.00	NA	NA	1.00	1.00
6	SelfC	=~	P6	0.92	0.06	16.05	0.00	0.81	1.03
7	Pain	=~	P7	1.00	0.00	NA	NA	1.00	1.00
8	Pain	=~	P8	1.08	0.05	21.67	0.00	0.99	1.18
9	Cogni	=~	P9	1.00	0.00	NA	NA	1.00	1.00
10	Cogni	=~	P10	0.99	0.06	15.56	0.00	0.87	1.12
11	Interp	=~	P11	1.00	0.00	NA	NA	1.00	1.00
12	Interp	=~	P12	0.98	0.10	10.14	0.00	0.79	1.17
13	Vision	=~	P13	1.00	0.00	NA	NA	1.00	1.00
14	Vision	=~	P14	1.02	0.09	11.18	0.00	0.84	1.20
15	Energ	=~	P15	1.00	0.00	NA	NA	1.00	1.00
16	Energ	=~	P16	0.76	0.06	12.87	0.00	0.64	0.87
17	Affect	=~	P17	1.00	0.00	NA	NA	1.00	1.00
18	Affect	=~	P18	0.91	0.04	22.12	0.00	0.83	0.99
19	Disab	=~	GHealth	1.00	0.00	NA	NA	1.00	1.00
20	Disab	=~	Movi	1.21	0.07	16.88	0.00	1.07	1.36
21	Disab	=~	SelfC	0.55	0.07	7.74	0.00	0.41	0.70
22	Disab	=~	Pain	1.24	0.09	13.75	0.00	1.06	1.41
23	Disab	=~	Cogni	1.04	0.08	13.33	0.00	0.89	1.19
24	Disab	=~	Interp	0.87	0.10	9.10	0.00	0.68	1.06
25	Disab	=~	Vision	0.81	0.08	9.68	0.00	0.65	0.97
26	Disab	=~	Energ	1.30	0.10	12.45	0.00	1.09	1.50
27	Disab	=~	Affect	1.34	0.08	17.38	0.00	1.19	1.49
28	P1	~~	P1	0.35	0.02	16.61	0.00	0.31	0.39
29	P2	~~	P2	0.42	0.04	10.88	0.00	0.34	0.50
30	P3	~~	P3	0.36	0.04	9.73	0.00	0.29	0.44
31	P4	~~	P4	1.51	0.08	18.35	0.00	1.35	1.67
32	P5	~~	P5	0.06	0.01	4.14	0.00	0.03	0.08
33	P6	~~	P6	0.10	0.02	4.44	0.00	0.05	0.14
34	P7	~~	P7	0.36	0.05	6.85	0.00	0.26	0.47
35	P8	~~	P8	0.23	0.04	6.46	0.00	0.16	0.30
36	P9	~~	P9	0.54	0.05	10.62	0.00	0.44	0.64
37	P10	~~	P10	0.38	0.04	9.67	0.00	0.30	0.46

38	P11	~~	P11	0.45	0.08	5.91	0.00	0.30	0.60
39	P12	~~	P12	0.57	0.06	8.97	0.00	0.45	0.70
40	P13	~~	P13	0.58	0.07	8.32	0.00	0.44	0.71
41	P14	~~	P14	0.54	0.07	8.26	0.00	0.41	0.67
42	P15	~~	P15	0.49	0.06	8.32	0.00	0.37	0.60
43	P16	~~	P16	0.57	0.05	10.79	0.00	0.47	0.67
44	P17	~~	P17	0.33	0.05	7.27	0.00	0.24	0.42
45	P18	~~	P18	0.57	0.04	12.96	0.00	0.48	0.66
46	GHealth	~~	GHealth	0.01	0.01	0.55	0.58	-0.01	0.03
47	Movi	~~	Movi	0.10	0.02	4.19	0.00	0.05	0.15
48	SelfC	~~	SelfC	0.23	0.04	5.62	0.00	0.15	0.31
49	Pain	~~	Pain	0.39	0.05	7.98	0.00	0.30	0.49
50	Cogni	~~	Cogni	0.28	0.04	7.98	0.00	0.21	0.35
51	Interp	~~	Interp	0.17	0.03	5.71	0.00	0.11	0.23
52	Vision	~~	Vision	0.40	0.05	7.23	0.00	0.29	0.50
53	Energ	~~	Energ	0.55	0.07	7.94	0.00	0.42	0.69
54	Affect	~~	Affect	0.40	0.04	9.04	0.00	0.31	0.48
55	Disab	~~	Disab	0.28	0.03	8.77	0.00	0.22	0.35
56	P1	~1		1.36	0.03	48.77	0.00	1.31	1.42
57	P2	~1		0.72	0.03	22.77	0.00	0.66	0.78
58	P3	~1		0.49	0.03	16.25	0.00	0.43	0.55
59	P4	~1		1.41	0.05	29.92	0.00	1.32	1.50
60	P5	~1		0.16	0.02	7.40	0.00	0.12	0.20
61	P6	~1		0.17	0.02	7.58	0.00	0.12	0.21
62	P7	~1		1.22	0.04	34.48	0.00	1.15	1.29
63	P8	~1		1.20	0.04	33.56	0.00	1.13	1.27
64	P9	~1		0.95	0.03	27.57	0.00	0.88	1.02
65	P10	~1		0.58	0.03	17.93	0.00	0.52	0.64
66	P11	~1		0.43	0.03	13.41	0.00	0.37	0.49
67	P12	~1		0.62	0.03	18.12	0.00	0.55	0.68
68	P13	~1		0.64	0.03	18.66	0.00	0.57	0.70
69	P14	~1		0.63	0.03	18.77	0.00	0.57	0.70
70	P15	~1		1.08	0.04	25.93	0.00	1.00	1.16
71	P16	~1		0.96	0.03	27.79	0.00	0.89	1.02
72	P17	~1		1.08	0.04	28.66	0.00	1.01	1.16
73	P18	~1		1.30	0.04	34.56	0.00	1.23	1.38
74	GHealth	~1		0.00	0.00	NA	NA	0.00	0.00
75	Movi	~1		0.00	0.00	NA	NA	0.00	0.00
76	SelfC	~1		0.00	0.00	NA	NA	0.00	0.00
77	Pain	~1		0.00	0.00	NA	NA	0.00	0.00
78	Cogni	~1		0.00	0.00	NA	NA	0.00	0.00
79	Interp	~1		0.00	0.00	NA	NA	0.00	0.00
80	Vision	~1		0.00	0.00	NA	NA	0.00	0.00
81	Energ	~1		0	0.00	NA	NA	0.00	0.00
82	Affect	~1		0	0.00	NA	NA	0.00	0.00
83	Disab	~1		0	0.00	NA	NA	0.00	0.00

lhs: left hand side of the operator/ rhs: right hand side of the operator/ op: operator/ est: estimate/ se: standard error

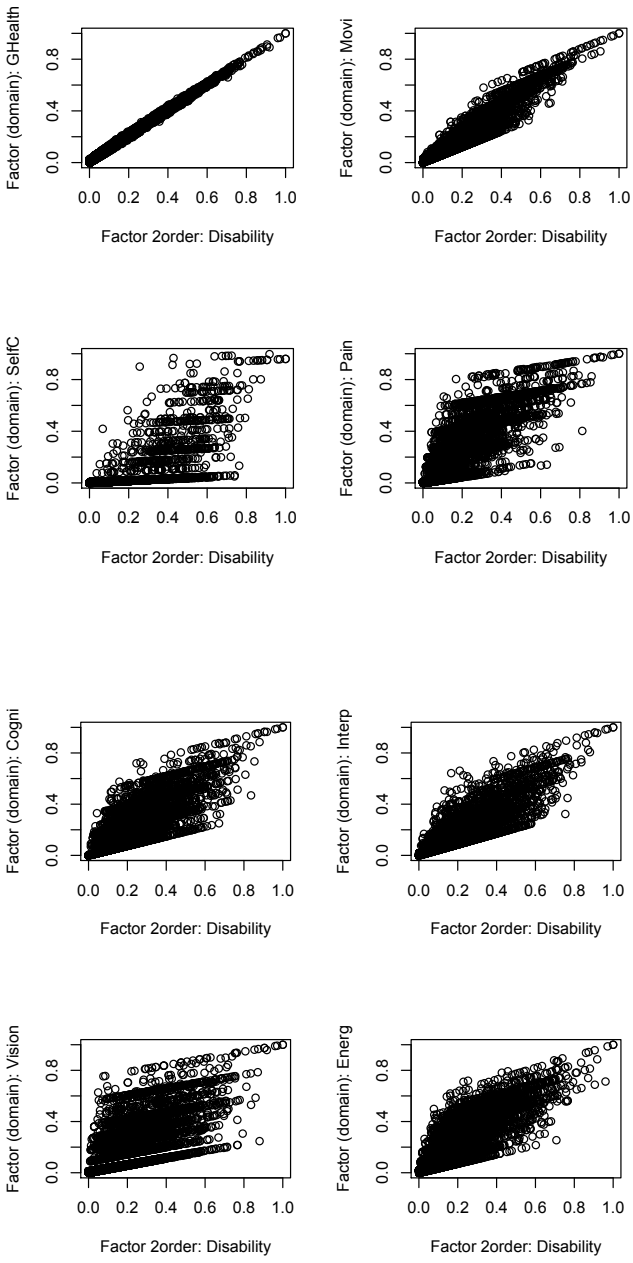
Disab: disability/ Movi: mobility domain/ SelfC: self-care domain/ Pain: pain and discomfort domain/ Cgni: cognition domain/ Interp: interpersonal activities domain/ Vision: vision domain/ Energ: sleep and energy domain/ Affect: affect domain/ GHealth: overall health domain/ P1-P18: represent the 18 items of the questionnaire.

Table S2.8. Correlation matrix between factor scores from the World Health Survey (HSD-WHS) questionnaire (n=5,293), using Confirmatory Factor Analysis, assuming continuous variables but including sample weights. Chilean National Health Survey 2009-2010.

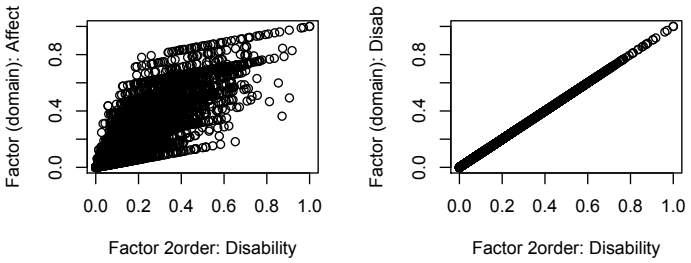
	GHealth	Movi	SelfC	Pain	Cogni	Interp	Vision	Energ	Affect	Disab
GHealth	1.00									
Movi	0.94	1.00								
SelfC	0.62	0.63	1.00							
Pain	0.81	0.75	0.42	1.00						
Cogni	0.81	0.75	0.50	0.60	1.00					
Interp	0.84	0.80	0.56	0.62	0.74	1.00				
Vision	0.66	0.61	0.42	0.48	0.55	0.53	1.00			
Energ	0.82	0.74	0.44	0.66	0.66	0.68	0.53	1.00		
Affect	0.81	0.69	0.40	0.64	0.64	0.67	0.50	0.74	1.00	
Disab	1.00	0.95	0.63	0.81	0.82	0.85	0.66	0.82	0.81	1.00

Disab: disability/ Movi: mobility domain/ SelfC: self-care domain/ Pain: pain and discomfort domain/ Cgni: cognition domain/ Interp: interpersonal activities domain/ Vision: vision domain/ Energ: sleep and energy domain/ Affect: affect domain/ GHealth: overall health domain.

Figure S2.5. Scatterplot of correlation between factor scores from different domains and second order factor score of disability. World Health Survey (HSD-WHS) questionnaire (n=5,293), using Confirmatory Factor Analysis, assuming continuous variables but including sample weights. Chilean National Health Survey 2009-2010.



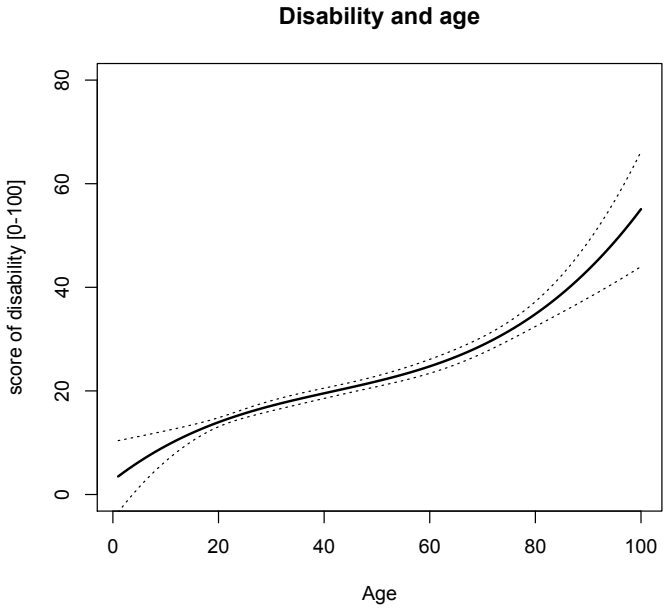
Disab: disability/ Movi: mobility domain/ SelfC: self-care domain/ Pain: pain and discomfort domain/ Cgni: cognition domain/ Interp: interpersonal activities domain/ Vision: vision domain/ Energ: sleep and energy domain/ Affect: affect domain/ GHealth: overall health domain.



Disab: disability/ Movi: mobility domain/ SelfC: self-care domain/ Pain: pain and discomfort domain/ Cgni: cognition domain/ Interp: interpersonal activities domain/ Vision: vision domain/ Energ: sleep and energy domain/ Affect: affect domain/ GHealth: overall health domain.

Figure S2.6. Predicted values of Disability (panel A) and Domains of Disability (panel B) according to Age. Chilean National Health Survey 2009-2010. (n=5,293), using Confirmatory Factor Analysis, assuming continuous variables but including sample weights.

Panel A



Panel B

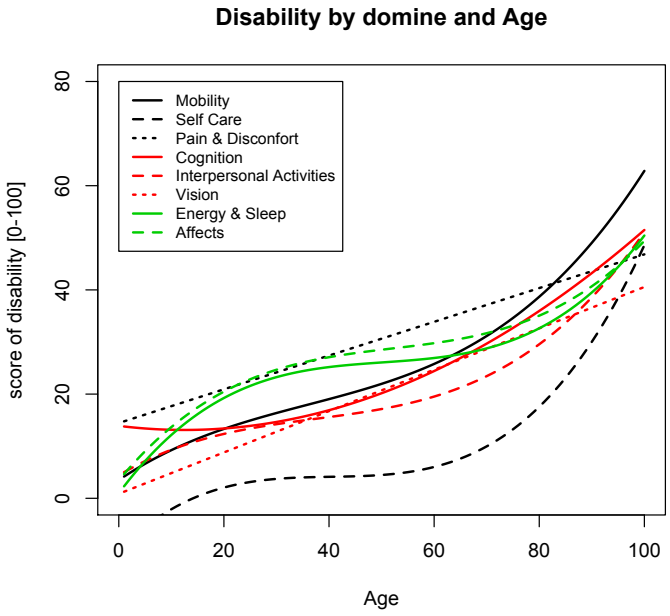


Table S2.9. Results of multivariate regression models on Disability and Domains of Disability. Chilean National Health Survey 2009-2010. (n=4,600)
Disability score and its Domains were calculated using a Confirmatory Factor Analysis, assuming continuous variables but including sample weights.

	Disability			Overall Health			Mobility		
	Coeff	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**
Intercept	10.3	7.6	13.0	10.6	7.9	13.3	8.1	5.0	11.1
Sex (females)	4.3	3.1	5.5	4.4	3.1	5.6	4.4	3.1	5.7
Age (10 years)	2.0	1.6	2.4	2.0	1.6	2.4	2.4	2.0	2.8
Education (< 8 years)	0.0	-	-	0.0	-	-	0.0	-	-
Education (8-12 years)	-4.0	-5.9	-2.1	-4.0	-5.9	-2.1	-4.6	-6.8	-2.5
Education (>12 years)	-6.2	-8.3	-4.1	-6.4	-8.5	-4.3	-7.0	-9.3	-4.7
Depression	14.7	12.8	16.6	14.7	12.8	16.6	13.9	11.9	16.0
Hypertension	2.1	0.3	3.9	2.1	0.3	3.9	3.1	1.1	5.0
Diabetes	5.4	2.3	8.5	5.5	2.4	8.6	6.3	2.7	9.9

	Self Care			Pain & Discomfort			Cognition		
	Coeff	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**
Intercept	0.0	-2.6	2.5	16.5	12.1	21.0	10.8	7.1	14.4
Sex (females)	0.3	-1.0	1.5	6.8	4.7	8.8	3.7	2.1	5.2
Age (10 years)	1.4	1.0	1.7	1.9	1.2	2.7	2.1	1.6	2.7
Education (< 8 years)	0.0	-	-	0.0	-	-	0.0	-	-
Education (8-12 years)	-2.1	-4.1	-0.1	-4.3	-7.2	-1.5	-5.8	-8.3	-3.4
Education (>12 years)	-3.6	-5.5	-1.8	-6.7	-10.1	-3.2	-8.1	-10.8	-5.4
Depression	4.2	2.5	5.8	14.9	12.0	17.7	13.6	11.0	16.3
Hypertension	0.2	-1.4	1.8	3.2	0.4	6.1	1.0	-1.2	3.2
Diabetes	4.0	-0.1	8.1	5.1	0.8	9.4	4.4	0.7	8.2

	Interpersonal activities			Vision			Spleep & Energy		
	Coeff	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**	**Coeff**	**LCI**	**UCI**
Intercept	9.3	6.2	12.5	3.4	0.0	6.8	12.4	8.7	16.1
Sex (females)	1.6	0.1	3.0	5.6	4.1	7.2	5.0	3.2	6.9
Age (10 years)	1.5	1.0	2.0	3.0	2.5	3.5	1.6	1.0	2.2
Education (< 8 years)	0.0	-	-	0.0	-	-	0.0	-	-
Education (8-12 years)	-3.8	-5.9	-1.6	-4.4	-6.8	-1.9	-1.2	-3.5	1.1
Education (>12 years)	-5.9	-8.2	-3.7	-7.5	-10.1	-5.0	-0.8	-3.7	2.0
Depression	14.8	12.6	17.0	10.7	8.3	13.1	18.7	16.2	21.3
Hypertension	1.8	-0.2	3.7	2.2	-0.1	4.5	0.5	-2.0	2.9
Diabetes	4.6	1.2	8.0	4.3	-0.1	8.6	4.3	0.9	7.6

	Affect		
	Coeff	**LCI**	**UCI**
Intercept	14.1	10.6	17.7
Sex (females)	5.5	3.7	7.2
Age (10 years)	1.6	1.0	2.2
Education (< 8 years)	0.0	-	-
Education (8-12 years)	-2.3	-4.7	0.2
Education (>12 years)	-2.9	-5.7	-0.1
Depression	24.8	22.3	27.3
Hypertension	1.0	-1.4	3.4
Diabetes	5.0	1.4	8.6

Coeff: regression coefficient/ LCI: lower confidence interval/ UCI: upper confidence interval

Figure S2.7. Distribution of the disability score in the general population and with three health conditions. Chilean National Health Survey 2009/2010. (n=4,600)

Disability score and its Domains were calculated using a Confirmatory Factor Analysis, assuming continuous variables but including sample weights.

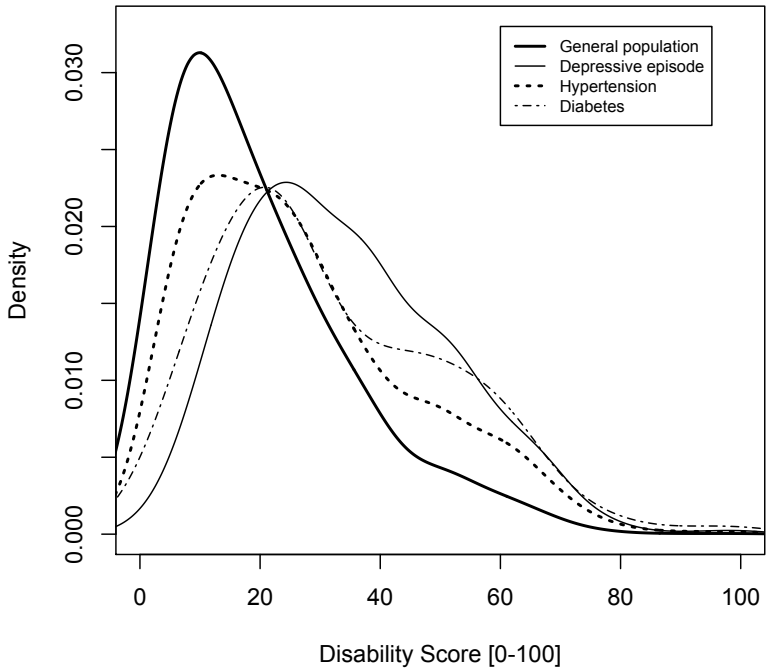
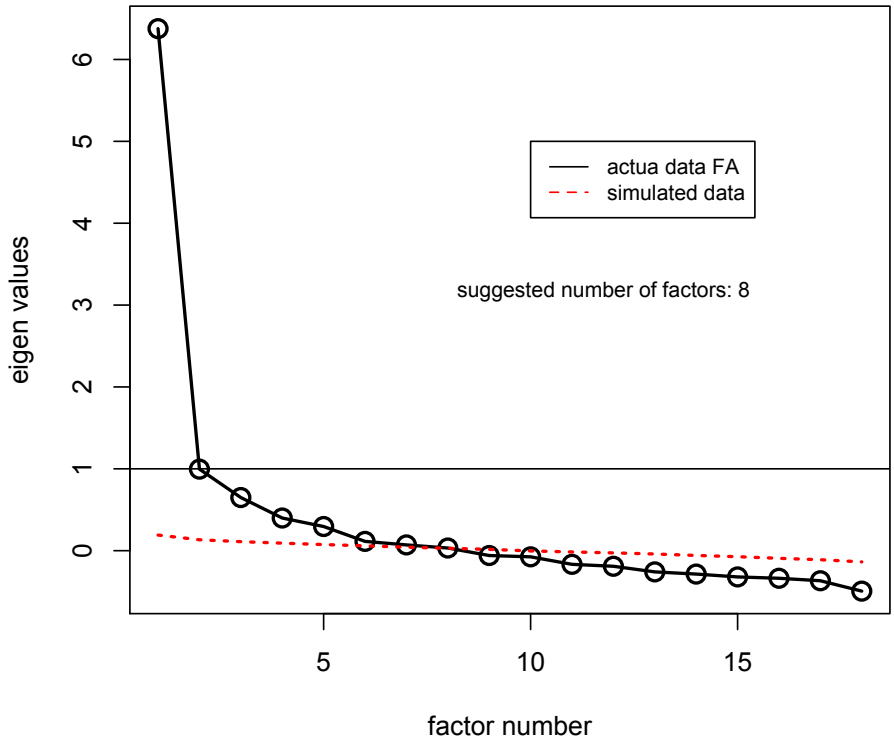


Figure S2.8. Parallel Analysis for choosing the number of factors to be explored. Chilean National Health Survey 2009-2010.

Sensitivity analysis using half of the sample with completed data



Method: weighted least square
N: 2,300

Table S2.10. Exploratory Factor Analysis: Factor loadings of each item with the eight subdomains of the Health State Description questionnaire of the World Health Survey (HSD-WHS). Chilean National Health Survey 2009-2010.
Sensitivity analysis using half of the sample with completed data

Item	Overall Health + Mobility	Self-Care	Pain & Discomfort	Cognition	Interpersonal activities	Vision	Sleep & Energy	Affect	Uniqueness
1	**0,43**	-	-	-	-	-	-	*0.27*	44%
2	**0,61**	-	-	-	-	-	-	-	23%
3	**0,76**	-	-	-	-	-	-	-	12%
4	**0,49**	-	-	-	-	-	-	-	55%
5	-	**0,93**	-	-	-	-	-	-	3%
6	-	**0,86**	-	-	-	-	-	-	4%
7	-	-	**0,94**	-	-	-	-	-	11%
8	-	-	**0,89**	-	-	-	-	-	8%
9	-	-	-	**0,83**	-	-	-	-	26%
10	-	-	-	**0,81**	-	-	-	-	22%
11	-	-	-	-	**0,63**	-	-	-	31%
12	-	-	-	-	**0,62**	-	-	-	41%
13	-	-	-	-	-	**0,78**	-	-	36%
14	-	-	-	-	-	**0,80**	-	-	36%
15	-	-	-	-	-	-	**0,73**	-	38%
16	-	-	-	-	-	-	**0,72**	-	36%
17	-	-	-	-	-	-	-	**0,74**	24%
18	-	-	-	-	-	-	-	**0,70**	30%

Factor loadings lower than 0.2 are omitted.
Larger factor loadings are marked with bold. Italic numbers show cross-loadings.
n 2,300/ method: weighted least square/ rotation: oblimin.

Table S2.11.Goodness of fit of a confirmatory factor analysis for the Health State Description questionnaire of the World Health Survey (HSD-WHS). Chilean National Health Survey 2009/2010. (n=2,300).
Sensitivity analysis using half of the sample with completed data

	Respecting categorical nature of variables			Using sample weight from the survey		
	Unidimentional model	model suggested by EFA	Theoretical model	Unidimentional model	model suggested by EFA	Theoretical model
χ^2	5804.1	1121.5	1157.4	6217.8	1492.5	1589.8
Df	135	127	126	135	127	126
p value	0	0	0	0	0	0
relative χ^2	43.0	8.8	9.2	46.1	11.8	12.6
TLI	0.93	0.99	0.99	0.60	0.90	0.90
CFI	0.94	0.99	0.99	0.65	0.92	0.92
GFI	0.96	0.99	0.99	0.82	0.95	0.95
RMSEA	0.135	0.058	0.060	0.140	0.068	0.071
RMSEA LCI	0.132	0.055	0.057	0.137	0.065	0.068
RMSA UCI	0.138	0.062	0.063	0.143	0.072	0.074
SRMR	0.108	0.059	0.059	0.086	0.051	0.052

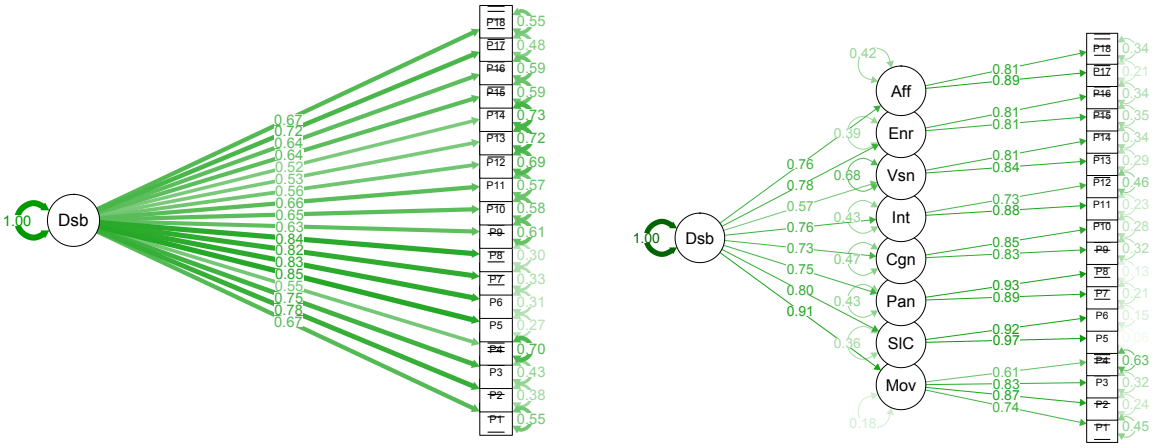
EFA: Exploratory factor analysis / χ^2: chi-square / Df: degree of freedom / TLI: Tucker-Lewis index / CFI: comparative fir index / GFI: goodness of fit index / RMSEA: root mean square error of approximation / SRMR: standardized root mean square residual / LCI: lower 90% confidence interval / UCI: upper 90% confidence interval

Figure S2.9. Path-diagram of the Health State Description of the World Health Survey (HSD-WHS). Chilean National Health Survey 2009/2010. (n=2,300)
Sensitivity analysis using half of the sample with completed data

Figure S.2.9.1 Path-diagrams considering categorical nature of data

A. Unidimentional model by EFA

B. Model according to the structure suggested



C. Model according to the structure suggested by the theoretical framework

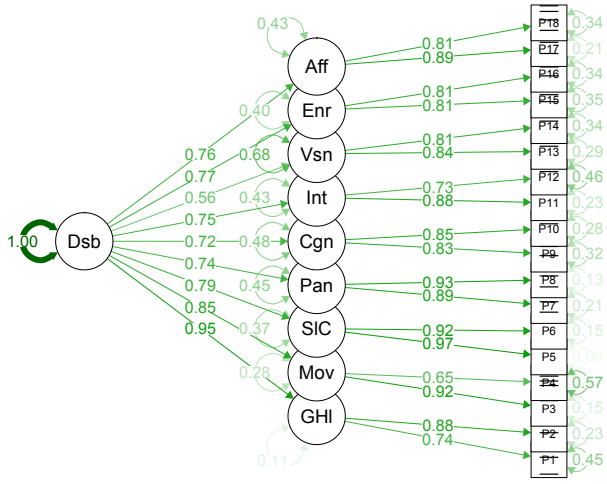
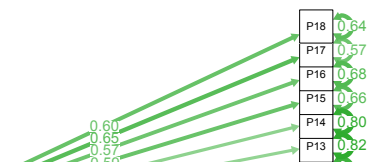
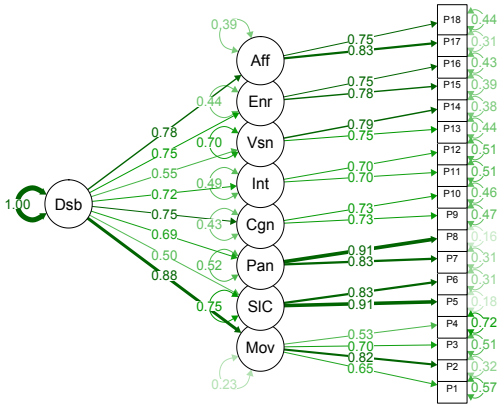


Figure S.2.9.2 Path-diagrams assuming continuous variables but considering sample weights



D. Unidimentional model by EFA

E. Model according to the structure suggested



F. Model according to the structure suggested by the theoretical framework

Considering categorical data

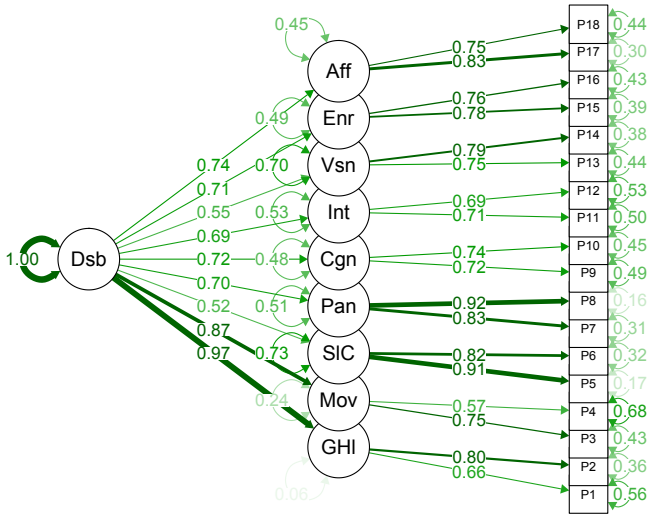


Table S2.12. Reliability of the Health State Description questionnaire of the World Health Survey (HSD-WHS), and description of factors scores. Chilean National Health Survey 2009/2010. (n=4,600)

Sensitivity analysis using the sample with completed data

Domine	std.alpha	mean	sd	median	Q1	Q3	Minnimum	Maximum
Overall Health	0.70	32.2	14.9	16.9	9.0	29.1	0.0	100.0
Mobility	0.64	32.3	15.7	14.5	7.5	26.7	0.0	100.0
Self Care	0.87	32.2	15.9	1.2	0.6	2.1	0.0	100.0
Pain & Discomfort	0.87	32.2	19.0	23.3	7.6	43.6	0.0	100.0
Cognition	0.75	30.0	17.2	13.2	4.3	28.3	0.0	100.0
Interpersonal	0.66	30.8	15.7	10.6	4.5	22.2	0.0	100.0
Vision	0.71	27.1	17.6	8.3	3.0	26.4	0.0	100.0
Sleep & Energy	0.73	31.9	17.1	19.8	6.9	37.7	0.0	100.0
Affects	0.79	33.3	18.0	22.1	8.8	39.6	0.0	100.0
Disability	-	32.0	14.7	16.7	9.0	28.7	0.0	100.0
Whole instrument	0.90	-	-	-	-	-	-	-

stda.alpha: standardised alpha/ sd: standard deviation / Q1: quantile 25% / Q3: quantile 75%

Supplementary material chapter 3

Contents

Additional methods S3.1. Methods for the analysis by subdomain

For the analysis by subdomains of disability, we estimated the burden of disease on each one of the disability scores by subdomain, following the same procedure described for overall disability. Additionally, we also wanted to explore the weight of each subdomain in the total disability and comparing their distributions across diseases. To do this, we calculated the fraction of disability attributable to each subdomain, contrasting two scenarios: the actual scenario which corresponds to the observed disability (D); and a counterfactual scenario in which the disability due to a certain subdomain is assumed 0 (D'). The assumption was implemented imputing the answer 'none difficulty or problem' in the items corresponding to the subdomain and recalculating the latent variable of disability. Consequently, the disability attributable to a certain domain can be expressed in absolute terms as D – D', as well as a fraction: (D-D')/D. The fraction of disability attributable to each subdomain was calculated for general population and for people with each health condition under study. To ensure comparability across individuals with different diseases, we calculated the predicted values of D and D' using the models presented for overall disability, setting the covariables in the population means.

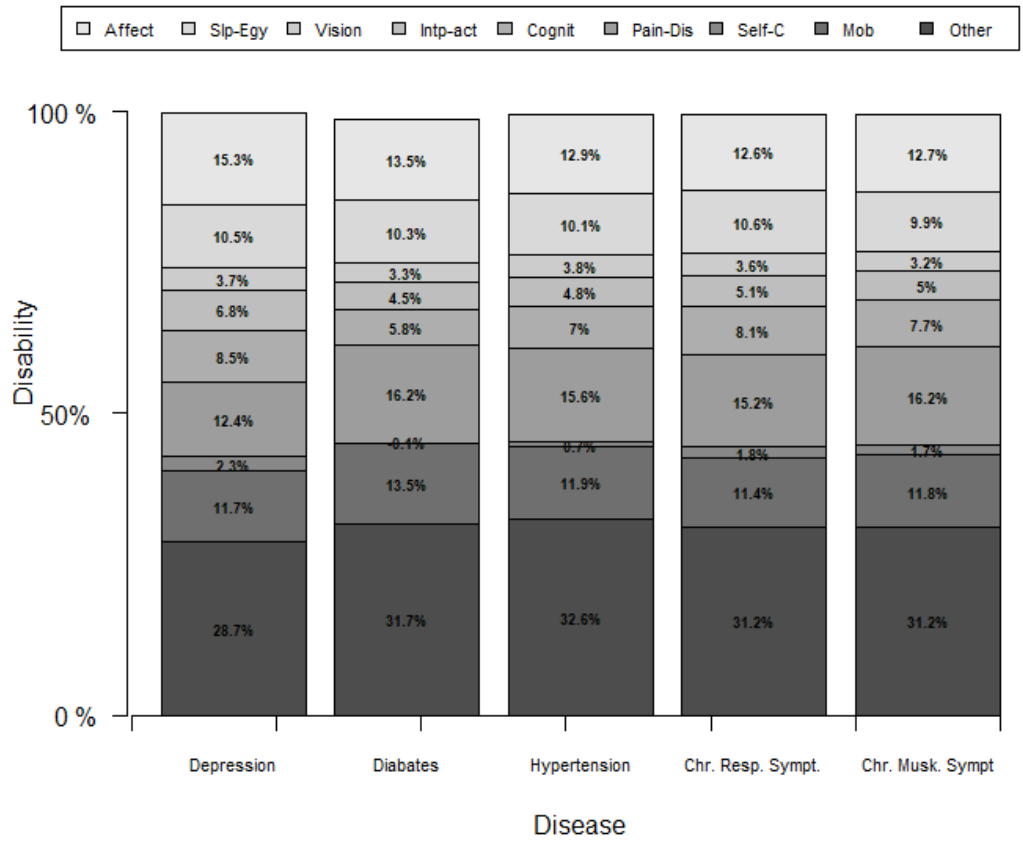
Table S3.1**.** Description of the sample, Chilean National Health Survey 2009-2019 (n=4.447).

	n	%*	CI95%
Sex			
female	2675	51.9	[49.3 - 54.4]
Age			
15-29	973	30.4	[28.0 - 32.8]
30-59	2335	52.1	[49.6 - 54.6]
60+	1139	17.5	[15.9 - 19.2]
Education			
> 12 yrs	832	23.7	[21.4 - 25.9]
9 - 12 yrs	2446	57.7	[55.2 - 60.2]
<8 yrs	1169	18.6	[16.9 - 20.3]
Morbidity			
Depressive episode	717	17.4	[15.5 - 19.3]
Diabetes	363	6.7	[5.6 - 7.9]
Hypertension	1488	27.4	[25.3 - 29.5]
Chronic. Respiratory symptoms	370	9	[7.5 - 10.6]
Chronic. Musculoskeletal symptoms	1601	31.1	[28.9 - 33.3]

**Percentages are weighted using the inverse of the probability of selection in sample.*

The most prevalent conditions were chronic musculoskeletal symptoms and hypertension, affecting each one up to a third of the population older than 15 years, while other diseases such as diabetes was present in less than 10%. The prevalence of hypertension increases steadily from middle ages, while chronic musculoskeletal symptoms starts to rise at an earlier age, keeping stable after the forties around 40%. The prevalence of depressive episode reaches a peak between forties and fifties, but then decreases progressively. Chronic respiratory symptoms shows a stable prevalence across ages, while diabetes reproduces the pattern of hypertension but with lower magnitudes.

Figure S3.1**.** Attributable fraction of different sub-domains of disability for five selected health conditions. Chilean National Health Survey 2009-2019 (n=4,447).



Chr. Resp. Sympt. : chronic respiratory symptoms/ Chr. Musk. Sympt: chronic musculoskeletal symptoms.
Subdomain of disability: Affect, Slp-Egy (sleep and energy), Vision, Intp-act (interpersonal activities), Cognit (cognition), Pain-Dis (pain and discomfort), Self-C (self-care), Mob (mobility), Other (overall health)

Sub analysis 1: results for men

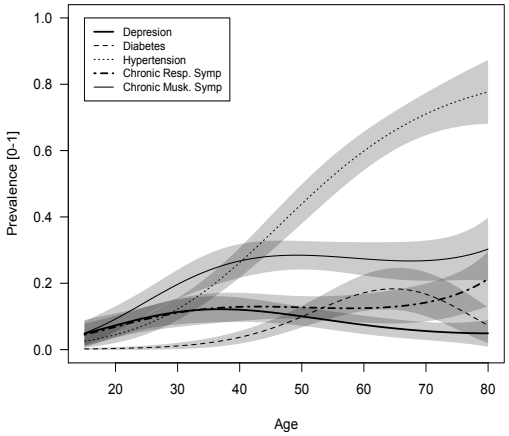
Table S3.2. Attributable disability [0 – 100], and burden of disease [%] on disability for different diseases according to different models in **men**. Chilean National Health Survey 2009-2010 (n=1,772).

	Univariate					
	D_A (disability attributable)		Burden (population approach)		Burden (individual approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depression	13.6	[9 to 18.2]	7.5	[4.8 - 10.7]	7.5	[6 - 9.2]
Diabetes	9.5	[3.4 to 15.6]	3.1	[1 - 5.5]	3.2	[2.4 - 4]
Hypertension	7.6	[4.9 to 10.3]	13.5	[8.9 - 18]	13.5	[11.8 - 15.3]
Chronic Resp Sympt.	8.5	[4.5 to 12.5]	5.9	[3 - 9.2]	6.0	[4.7 - 7.4]
Chromic Musk Sympt.	7	[4.8 to 9.2]	9.0	[6.1 - 12.1]	9.1	[7.9 - 10.4]
	Model 1					
	D_A (disability attributable)		Burden (population approach)		Burden (individual approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depression	11.9	[8 to 16]	6.5	[4.2 - 9.4]	6.6	[5.2 - 8]
Diabetes	5.4	[0 to 10.9]	1.8	[0 - 3.8]	1.8	[1.4 - 2.3]
Hypertension	2.3	[-0.3 to 5]	4.1	[-0.6 - 8.6]	4.1	[3.6 - 4.7]
Chronic Resp Sympt.	6.3	[2.2 to 10.4]	4.4	[1.6 - 7.4]	4.4	[3.5 - 5.5]
Chromic Musk Sympt.	4	[1.7 to 6.3]	5.2	[2.2 - 8.2]	5.2	[4.5 - 5.9]
	Model 2					
	D_A (disability attributable)		Burden (population approach)		Burden (individual approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depression	11.9	[7.6 to 16.2]	6.4	[4 - 9.4]	6.6	[5.2 - 7.9]
Diabetes	5.9	[0.6 to 11.3]	1.9	[0.2 - 3.9]	2.0	[1.5 - 2.5]
Hypertension	2.5	[-0.4 to 5.5]	4.4	[-0.8 - 9.5]	4.5	[3.9 - 5.1]
Chronic Resp Sympt.	6.1	[1.6 to 10.4]	4.1	[1.1 - 7.4]	4.2	[3.3 - 5.2]
Chromic Musk Sympt.	3.8	[1.1 to 6.5]	4.9	[1.5 - 8.4]	5.0	[4.2 - 5.8]

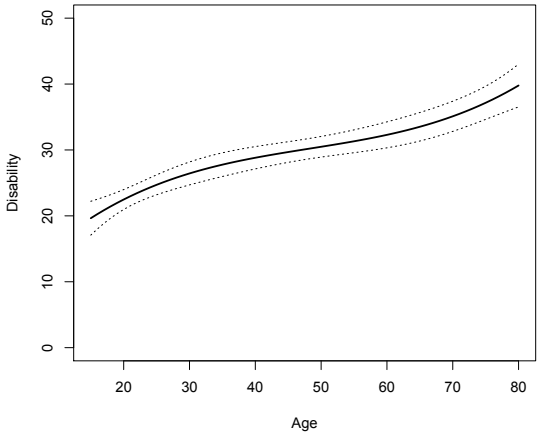
UI: uncertainty intervals (represent 2.5 and 97.5% quantile of resultant distribution)

Figure S3.2. Different inputs, and outputs related with the estimation of the burden Of disability attributable to diseases according to age, in **men**: Prevalence (A), disability (B), burden (C), weights of disability (D), weighted burden (E), and disability by subdomains (F). Chilean National Health Survey 2009-2010 (n=1,772).

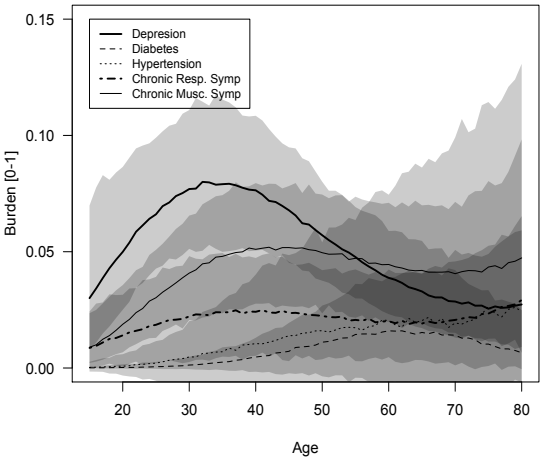
A. Prevalence of disease according to age



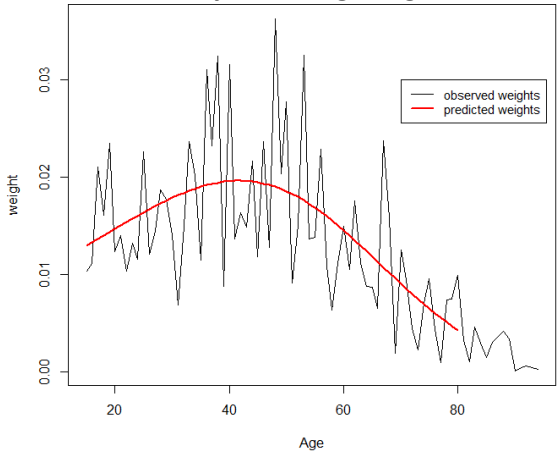
B. Predicted disability according to age



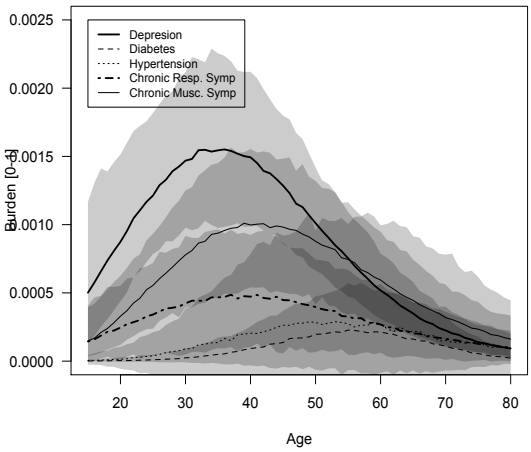
C. Burden of Disease according to age



D. Observed and predicted fraction of disability according to age



E. Burden of Disease according to age, weighted by the fraction of disability of each age



F. Predicted disability by subdomains according to age

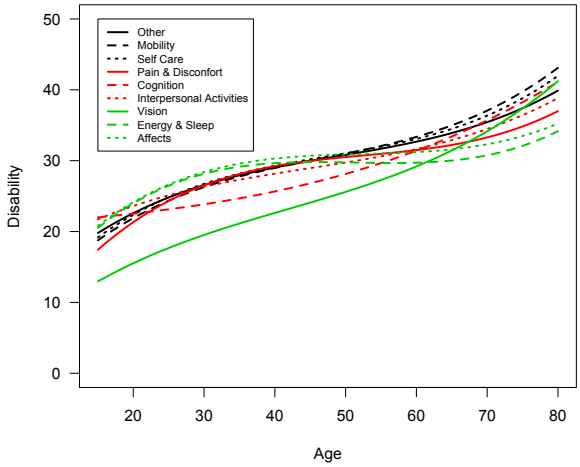
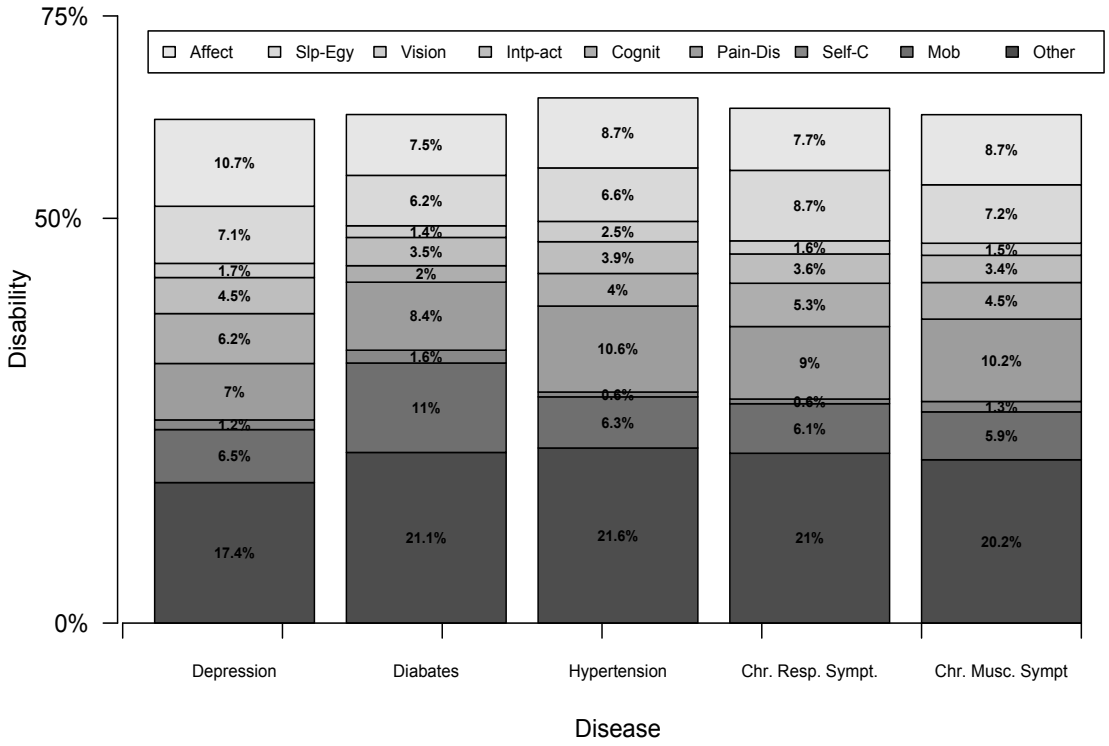
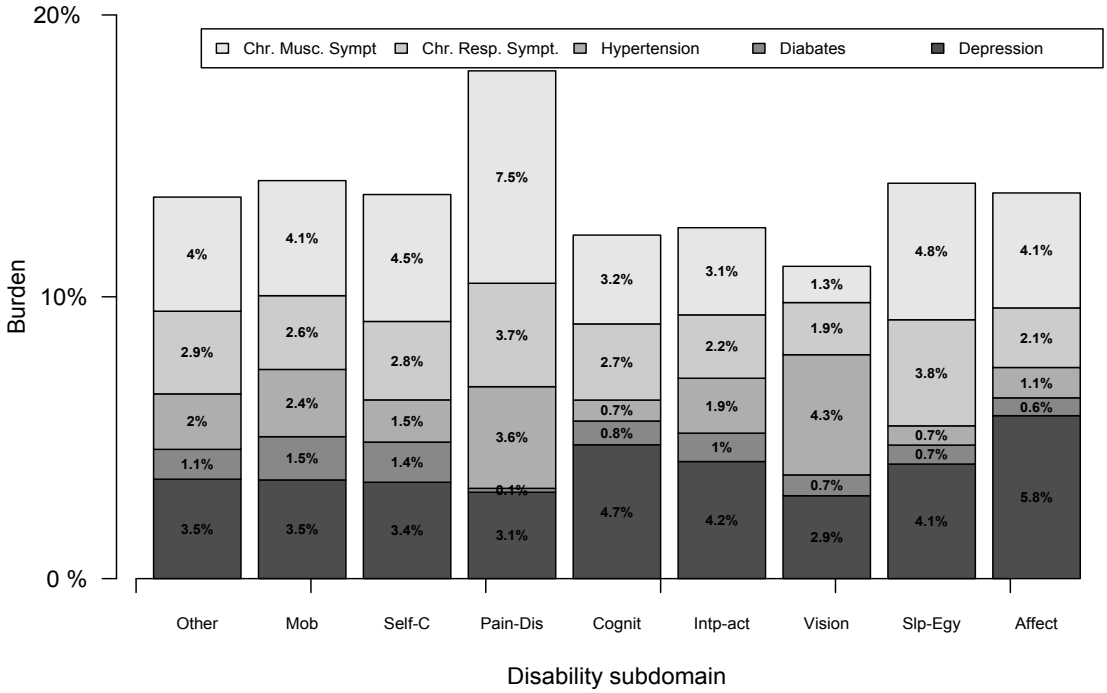


Figure S3.3 Attributable fraction of different subdomains in total disability for five health conditions, in **men**. Chilean National Health Survey 2009-2019 (n=1,772).



Chr. Resp. Sympt. : chronic respiratory symptoms/ Chr. Musk. Sympt: chronic musculoskeletal symptoms.
Subdomain of disability: Affect, Slp-Egy (sleep and energy), Vision, Intp-act (interpersonal activities), Cognit (cognition), Pain-Dis (pain and discomfort), Self-C (self-care), Mob (mobility), Other (overall health)

Figure S3.4. Burden of sub-domains of disability for five selected health conditions, in **men**. Chilean National Health Survey 2009-2019 (n=1,772).



Chr. Resp. Sympt. : chronic respiratory symptoms/ Chr. Musk. Sympt: chronic musculoskeletal symptoms.
Subdomain of disability: Affect, Slp-Egy (sleep and energy), Vision, Intp-act (interpersonal activities), Cognit (cognition), Pain-Dis (pain and discomfort), Self-C (self-care), Mob (mobility), Other (overall health)

Sub analysis 2: results for women

Table S3.3. Attributable disability [0 – 100], and burden of disease [%] on disability for different diseases according to different models in **women**. Chilean National Health Survey 2009-2010 (n=1,772).

	Univariate					
	D_A (disability attributable)		Burden (population approach)		Burden (individual approach)	
	Disability	**LCI**	**%**	**LCI**	**%**	**LCI**
Depression	15.9	[13.4 to 18.5]	16.8	[14 - 19.7]	16.8	[14.9 - 18.9]
Diabetes	12.6	[8.7 to 16.4]	4.1	[2.8 - 5.7]	4.1	[3.4 - 4.9]
Hypertension	9.9	[7.5 to 12.3]	10.7	[8.2 - 13.3]	10.8	[9.7 - 11.8]
Chronic Resp Sympt.	11.3	[6.4 to 16.2]	3.2	[1.8 - 4.8]	3.2	[2.7 - 3.8]
Chromic Musk Sympt.	9.6	[7.6 to 11.5]	16.1	[12.9 - 19.3]	16.1	[14.7 - 17.7]

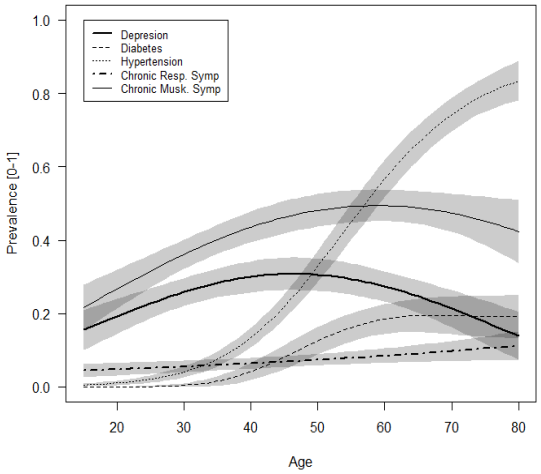
	Model 1					
	D_A (disability attributable)		Burden (population approach)		Burden (individual approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depression	13.8	[11.6 to 16.1]	14.6	[12.1 - 17.2]	14.6	[12.9 - 16.4]
Diabetes	4.4	[0.9 to 7.9]	1.4	[0.3 - 2.7]	1.4	[1.2 - 1.7]
Hypertension	1.7	[-0.7 to 3.9]	1.8	[-0.8 - 4.3]	1.8	[1.6 - 2]
Chronic Resp Sympt.	5.8	[2 to 9.5]	1.6	[0.6 - 2.8]	1.6	[1.3 - 1.9]
Chromic Musk Sympt.	5.1	[3.3 to 6.9]	8.6	[5.7 - 11.5]	8.6	[7.8 - 9.5]

	Model 2					
	D_A (disability attributable)		Burden (population approach)		Burden (individual approach)	
	Disability	**UI**	**%**	**UI**	**%**	**UI**
Depression	14.1	[11.5 to 16.7]	15.2	[12.2 - 18.2]	14.8	[13.1 - 16.7]
Diabetes	4.4	[0.9 to 8]	1.5	[0.3 - 2.7]	1.5	[1.2 - 1.7]
Hypertension	1.5	[-1.1 to 4]	1.6	[-1.2 - 4.4]	1.6	[1.5 - 1.8]
Chronic Resp Sympt.	5.5	[1.7 to 9.4]	1.6	[0.5 - 2.8]	1.6	[1.3 - 1.9]
Chromic Musk Sympt.	5.3	[3.1 to 7.4]	9.1	[5.5 - 12.6]	8.9	[8.1 - 9.8]

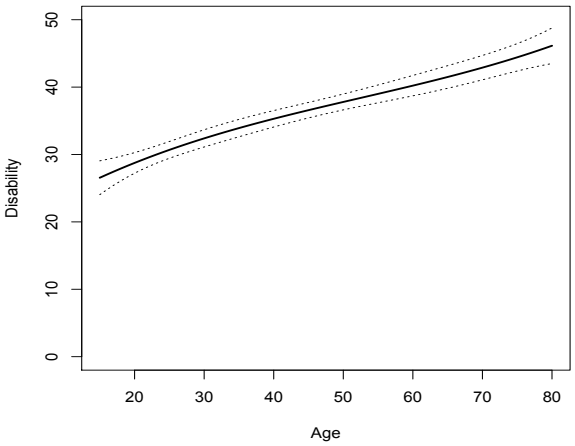
UI: uncertainty intervals (represent 2.5 and 97.5% quantile of resultant distribution)

Figure S3.5 Different inputs, and outputs related to the process of burden of disease on disability estimation according to age, in **women**: Prevalence (A), disability (B), burden (C), weights of disability (D), weighted burden (E), and disability by subdomains (F). Chilean National Health Survey 2009-2010 (n=2,675).

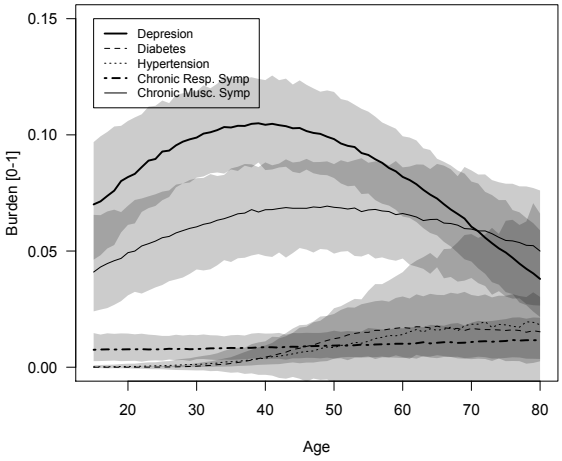
A. Prevalence of disease according to age



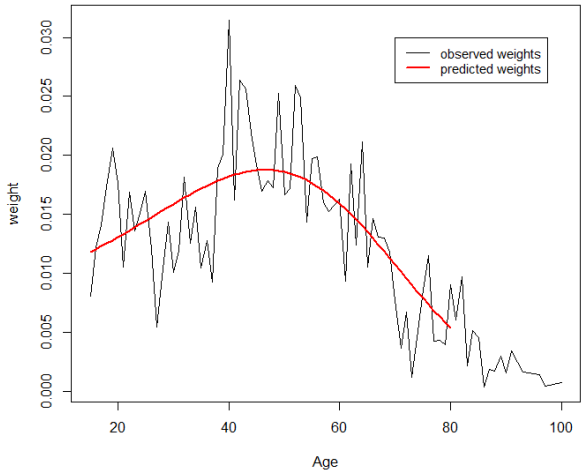
B. Predicted disability according to age



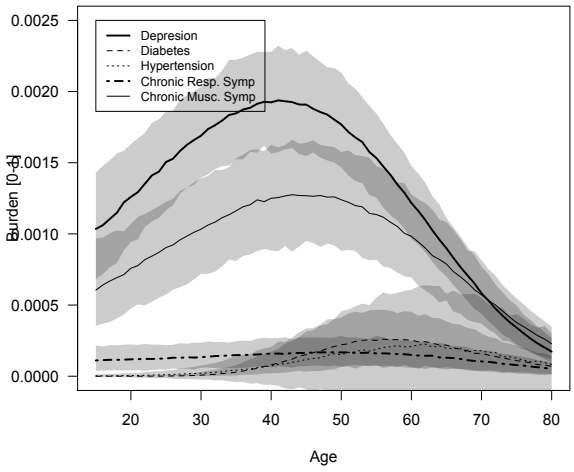
C. Burden of Disease according to age



D. Observed and predicted fraction of disability according to age



E. Burden of Disease according to age, weighted fraction of disability of each age



F. Predicted disability by the subdomains according to age

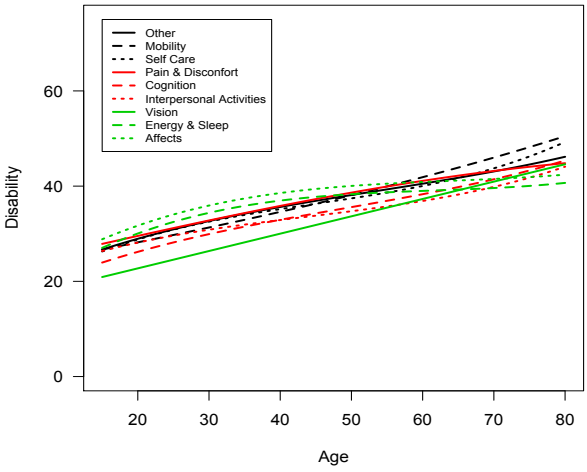
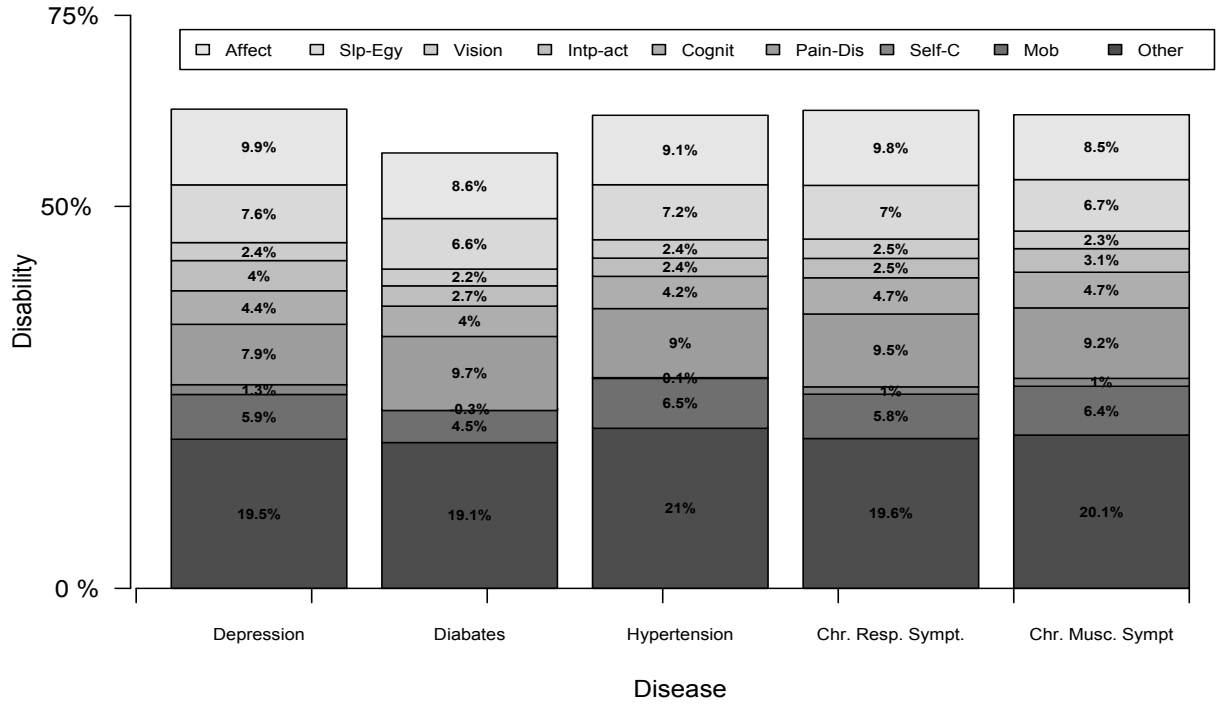
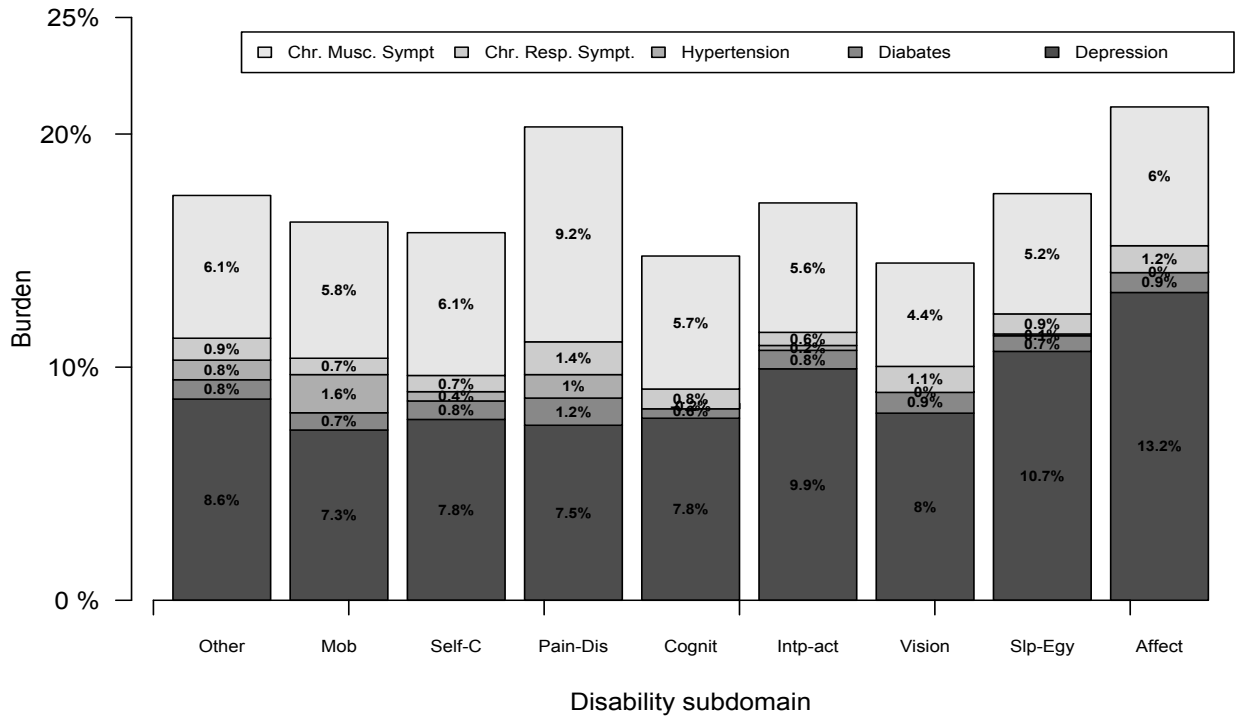


Figure S3.6. Attributable fraction of different sub-domains of disability for five selected health conditions, in **women**. Chilean National Health Survey 2009-2019 (n=2,675).



Chr. Resp. Sympt. : chronic respiratory symptoms/ Chr. Musk. Sympt: chronic musculoskeletal symptoms.
Subdomain of disability: Affect, Slp-Egy (sleep and energy), Vision, Intp-act (interpersonal activities), Cognit (cognition), Pain-Dis (pain and discomfort), Self-C (self-care), Mob (mobility), Other (overall health)

Figure S3.7. Burden of disease on subdomains of disability, for five selected health conditions, in **women**. Chilean National Health Survey 2009-2019 (n=2,675).



Chr. Resp. Sympt. : chronic respiratory symptoms/ Chr. Musk. Sympt: chronic musculoskeletal symptoms.
Subdomain of disability: Affect, Slp-Egy (sleep and energy), Vision, Intp-act (interpersonal activities), Cognit (cognition), Pain-Dis (pain and discomfort), Self-C (self-care), Mob (mobility), Other (overall health)

Main functions for R, used to estimate the burden of disability attributable to diseases

```
library(survey)

#Funtion to estimate Prevalence
# Input: Disease   (name of variables that contain the information about the presence or absence of the disease,
                   i.e. "DepreDSM")
# Input: Sex (name of the variable with Sex information)
# Input: Age (nams of the variable with Age information, non-centred and in x/10 scale)
# Input: design (svy design object)

Preval=function(Disease, Sex, Age, design) {

                   mod=design
                   a=svymean(as.formula(paste0("~",Disease)), mod)
                   SaP=c(coef(a), coef(a)-1.96*SE(a),  coef(a)+1.96*SE(a))

                   a=svyby(as.formula(paste0("~",Disease)), as.formula(paste0("~",Sex)), mod,
                   svymean, vartype="ci")
                   SaS=a[,-1]

                   mo=svyglm(as.formula(paste0(Disease, "~ ", Age, " + I(", Age,"^2) + I(", Age, "^3)")),
                   mod, family="binomial")
                   mo=step(mo)
                   da=data.frame(seq(1.5,8, 0.1));      names(da)=Age
                   SaE=predict(mo, da, type="response")
                   SaE=cbind(coef(SaE), coef(SaE)-1.96*SE(SaE), coef(SaE)+1.96*SE(SaE))

                   if (Disease==Sex) {SaES=matrix(rep(NA,3*nrow(da)*2), ncol=3)} else {
                   mo=svyglm(as.formula(paste0(Disease, "~ ", Age, " + I(", Age, "^2) + I(", Age, "^3) +
                   ", Sex, " + ", Sex, ":",Age)), mod, family="binomial")
                   mo=step(mo)
                   mo=update(mo,as.formula(paste0(" . ~ + ", Sex, " + ", Age)))
                   Ed= rep(seq(1.5,8, 0.1),2)
                   L= length(Ed)/2
                   Sexo=rep(c(0,1), each=L)
                   da=data.frame(cbind(Ed, Sexo)); names(da)=c(Age,Sex)
                   SaES=predict(mo,da, type="response")
                   SaES=cbind(coef(SaES), coef(SaES)-1.96*SE(SaES), coef(SaES)+1.96*SE(SaES)) }

                   return (list(Prev=SaP, Sex=SaS, Age=SaE, AgeS=list(M=SaES[1:L,],
                   F=SaES[(L+1):nrow(SaES),])))}


# Example about how to use the function on a vector of diseases

mod=svydesign(~1, weights=~weights, data=data)
Enf=c("DepreDSM", "DM",  "HTA", "SintResp", "SintMusq")
Sal=lapply(Enf, Preval, "Sexo", "Edad2", design=mod)
```

#Function to specify a regression model
Input: Outcome (name of the dependent variable of the model)
Input: Sex (name of the variable with Sex information)
Input: Age (name of the variable with Age information, not centred and in x/10 scale)
Input: Cova (vector with names of covariables, excluding Age and Sex)
Input: Step (TRUE or FALSE, backward selection model should be implemented?)
Input: CubAge (TRUE or FALSE, cuadratic + cubic term for Age should be added to the model?)
Input: interactAS (TRUE or FALSE, interaction term between Age and Sex should be added to the model?)
Input: interactWithAS (vector of covariables that should interact with sex and age)
Input: design (svy design object)

```
Regre=function(Outcome, Age, Sex, Cova, Step, CubAge, InteractAS, InteractWithAS, design) {

                mod=design
                 Covari=paste("+",Cova, collapse="")
                 Cub=paste0("+ I(",Age,"^2)", " + I(",Age,"^3)")
                 Int1=paste0(" + ", Sex, ":", Age)
                 Int2=paste0(Sex, ":", InteractWithAS, collapse=" + ");Int2=paste0(" + ",Int2)
                 Int3=paste0(Age, ":", InteractWithAS, collapse=" + ");          Int3=paste0(" + ",Int3)
                 Re=paste0(Outcome, "~ ", Sex, " + ", Age, Covari)

                if (CubAge==TRUE) {                  Re=paste0(Re, Cub) }
                if (InteractAS==TRUE) {              Re=paste0(Re, Int1) }
                 if(!is.null(InteractWithAS)){                  Re=paste0(Re, Int2, Int3) }
                 if (Step==TRUE) {modelo=svyglm(as.formula(Re), mod)
                                  modelo=step(modelo)
                                   Re=formula(modelo)}
                return(Re)}
```

Example about how to use the function for specify the regression model

```
mod=svydesign(~1, weights=~weights, data=data)
Regression=Regre(Outcome="Disability", Age="Age", Sex="Sex", Cova=c("Education1", "Education2",
"Depression", "Diabetes", "Hypertension", "RespiratorySymptoms", "MusculosqueletalSymptoms"),
Step=TRUE, CubAge=TRUE, InteractAS=TRUE, InteractWithAS=NULL, design=mod)
```

```
#Function to calculate means of a variable
# Input: Vari (Variable which mean want to be obtained)
# Input: design (svy design object)

PopMeans=function(Vari, design) { Sa=svymean(as.formula(paste0("~", Vari)), design, na.rm=TRUE)
                                  names(Sa)=NULL
                                  return(coef(Sa))}


# Example about how to use the function to calculate the mean on a set of variables

mod=svydesign(~1, weights=~weights, data=data)
Vari=c("Disability", "Age", "Sex", "Education1", "Education2", "Depression", "Diabetes")
Sal=sapply(VariAll, PopMeans, mod)
```

```
#Function to calculate the Burden (population approach)
# Input: Regression (a character object that specifies the regression model)
# Input: data (the database where variables are from)
# Input: SW (the name of the variable with the sample weights)


Burden=function(Regression, data, SW) {
                        mod=svydesign(~1, weights= as.formula(paste0("~",SW)), data=data)
                         modelo=svyglm(as.formula(Regression), mod)
                         VariAll=attributes(terms(modelo))$term.labels
                         toElimin= c(-grep(":", VariAll), -grep("I", VariAll))
                         if (length(toElimin)>0) {    VariAll=VariAll[c(-grep(":", VariAll), -grep("I",
                VariAll))]            }

                         if (length(VariAll)>1) {        a= lapply(dat[,VariAll], function(x)
                                                          length(table(x)))
                                                          a=which(!a==2)
                                                         Vari=VariAll[-a] } else {
                                                          if (!length(table(dat[,VariAll]))==2)
                                                          {stop("variable is not a factor")} else
                                                          {Vari=VariAll}}

                         #Da: Disability attributable to the factor (disease)
                        #D0: predicted disability in people without the factor (disease)
                        #D1: predicted disability in people with the factor (disease)
                        #P: the prevalence of the factor (disease)
                        n=0
                         Sa=list()
                        for (i in Vari) {
                                                n=n+1
                                                P=svymean(as.formula(paste0("~",i)), mod)

                                                 dat0=dat[dat[,i]==0,]
                                                 dat1=dat[dat[,i]==1,]

                                                 mod0= svydesign(~1,
                                                 weights=as.formula(paste0("~",SW)), data=dat0)
                                                 mod1= svydesign(~1,
                                                 weights=as.formula(paste0("~",SW)), data=dat1)

                                                 da0=sapply(VariAll, PopMeans, mod0)
                                                 da1=sapply(VariAll, PopMeans, mod1)
                                                 da1prima=da1;   da1prima[i]=0
                                                 D0=predict(modelo, data.frame(t(da0)))
                                                 D1=predict(modelo, data.frame(t(da1)))
                                                D1prima= predict(modelo, data.frame(t(da1prima)))

                                                mu=coef(P)
                                                Va=SE(P)^2
                                                 alpha=(((1-mu)/Va)- (1/mu))*mu^2; alpha=abs(alpha)
                                                 beta=alpha*((1/mu)-1); beta=abs(beta)
                                                 Prev=mapply(rbeta, n=10000, alpha, beta)

                                                mu=c(coef(D0), coef(D1), coef(D1prima))
                                                Va=c(SE(D0), SE(D1), SE(D1prima))
                                                D=mapply(rnorm, 10000, mu, Va)
```

```
                                        D=cbind(D, D[,2]-D[,3])
                                        colnames(D)=c("D0", "D1", "D1prima", "Da")

                                        Bu= (D[,"Da"] * Prev)  /  ( (Prev*D[,"D1"]) + ((1-Prev)*
D[,"D0"]) )
                                         Bu= quantile(Bu, c(0.5, 0.025, 0.975))

                                         Sa[[n]]=list(
                                                    Burden=Bu,
                                                    Prev=c(coef(P), coef(P)-1.96*SE(P),
                                        coef(P)+1.96*SE(P)),
                                                   DA= quantile(D[,"Da"], c(0.5, 0.025, 0.975)),
                                                    D0=c(coef(D0), coef(D0)-1.96*SE(D0),
                                                    coef(D0)+1.96*SE(D0)),
                                                    D1=c(coef(D1), coef(D1)-1.96*SE(D1),
                                                    coef(D1)+1.96*SE(D1))
                                                    )}

                                #Tidying up the outputs
                               Burden= t(sapply(Sa, '[[', "Burden"))
                               Prev= t(sapply(Sa, '[[', "Prev"))
                               DA= t(sapply(Sa, '[[', "DA"))
                                D0=t(sapply(Sa, '[[', "D0"))
                                D1= t(sapply(Sa, '[[', "D1"))
                                rownames(Burden)=rownames(Prev)=rownames(DA)=rownames(D0)=row
                                names(D1)=Vari
                                Sa=list(Burden=Burden, Prev=Prev, DA=DA, D0=D0, D1=D1)

                               return(Sa)}


#Example 1, how to use the function on a set of variables through univariable regression models
Cova=c("Sex", "Education1", "Education2", "Depression", "Diabetes", "Hypertension",
"RespiratorySymptoms", "MusculosqueletalSymptoms")
Regression=paste0("Disability ~ ", Cova)
Sa=lapply(Regression, Burden, data=dat, SW="factor_f1")
write.table(t(sapply(Sa, '[[', "Burden")), dec=",")
write.table(t(sapply(Sa, '[[', "Prev")), dec=",")
write.table(t(sapply(Sa, '[[', "DA")), dec=",")
write.table(t(sapply(Sa, '[[', "D0")), dec=",")
write.table(t(sapply(Sa, '[[', "D1")), dec=",")


#Example 2, how to use the function on a set of variables through adjusted models
#Multivivariate (model 1)
Regression1=Regre          (Outcome="Disability", Age="Age", Sex="Sex", Cova=c"Education1",
                           "Education2", "Depression", "Diabetes", "Hypertension",
                           "RespiratorySymptoms", "MusculosqueletalSymptoms"), Step=FALSE,
                           CubAge=FALSE, InteractAS=FALSE, InteractWithAS=NULL, design=mod)
Sa=Burden(Regression1, data=dat, SW="factor_f1")
write.table(Sa, dec=",")
```

```
#Function to calculate Burden (individual approach)
# Input: Regression (a character object that specifies the regression model)
# Input: data (the database where variables are)
# Input: WF (the name of the variable with the sample weights)

BurdenIndiv=function(Regression, data, SW) {
                    mod=svydesign(~1, weights= as.formula(paste0("~",SW)), data=data)
                     modelo=svyglm(as.formula(Regression), mod)
                     VariAll=attributes(terms(modelo))$term.labels
                     toElimin= c(-grep(":", VariAll), -grep("I", VariAll))
                     if (length(toElimin)>0) {     VariAll=VariAll[c(-grep(":", VariAll), -grep("I",
VariAll))]              }

                     if (length(VariAll)>1) {        a= lapply(dat[,VariAll], function(x)
                                                      length(table(x)))
                                                      a=which(!a==2)
                                                     Vari=VariAll[-a] } else {
                                                      if (!length(table(dat[,VariAll]))==2)
                                                      {stop("variable is not a factor")} else
                                                      {Vari=VariAll}}

                     #Da: Disability attributable to the factor (disease)
                    #D0: predicted disability in people without the factor (disease)
                    #D1: predicted disability in people with the factor (disease)
                    n=0
                     Sa=list()
                    for (i in Vari) {
                                          n=n+1

                                           dat0=dat[dat[,i]==0,]
                                           dat1=dat[dat[,i]==1,]
                                           dat1prima=dat1;              dat1prima[,i]=0

                                           D0=predict(modelo, dat0 )
                                           D1=predict(modelo, dat1)
                                          D1prima= predict(modelo, dat1prima)
                                          DA=D1-D1prima

                                           mod0= svydesign(~1,
                                           weights=as.formula(paste0("~",SW)), data=dat0)
                                           mod1= svydesign(~1,
                                           weights=as.formula(paste0("~",SW)), data=dat1)

                                          D0=svytotal(~D0, mod0)
                                          D1=svytotal(~D1, mod1)
                                           DA=svytotal(~DA, mod1)

                                          mu=c(coef(D0), coef(D1), coef(DA))
                                          Va=c(SE(D0), SE(D1), SE(DA))
                                          D=mapply(rnorm, 10000, mu, Va)

                                          colnames(D)=c("D0", "D1", "DA")

                                          Bu= D[,"DA"] / (D[,"D0"] + D[,"D1"])
                                           Bu= quantile(Bu, c(0.5, 0.025, 0.975))

                                           Sa[[n]]=list(
```

```
                                                    Burden=Bu,
                                                   DA= quantile(D[,"DA"], c(0.5, 0.025, 0.975)),
                                                    D0=c(coef(D0), coef(D0)-1.96*SE(D0),
                                                    coef(D0)+1.96*SE(D0)),
                                                    D1=c(coef(D1), coef(D1)-1.96*SE(D1),
                                                    coef(D1)+1.96*SE(D1))
                                                    )}

                                  #Tidying up the outputs
                                  Burden= t(sapply(Sa, '[[', "Burden"))
                                  Prev= t(sapply(Sa, '[[', "Prev"))
                                  DA= t(sapply(Sa, '[[', "DA"))
                                   D0=t(sapply(Sa, '[[', "D0"))
                                   D1= t(sapply(Sa, '[[', "D1"))
                                   rownames(Burden)=rownames(DA)=rownames(D0)=rownames(D1)=Vari
                                   Sa=list(Burden=Burden, DA=DA, D0=D0, D1=D1)

                                  return(Sa)}


#Example 1, how to use the function on a set of variables through univariable regression models
Cova=c("Sex", "Education1", "Education2", "Depression", "Diabetes", "Hypertension",
"RespiratorySymptoms", "MusculosqueletalSymptoms")
Regression=paste0("Disability ~ ", Cova)
Sa=lapply(Regression, BurdenIndiv, data=dat, SW="factor_f1")
write.table(t(sapply(Sa, '[[', "Burden")), dec=",")


#Example 2, how to use the function on a set of variables through adjusted models
#Multivivariate (model 1)
Regression1=Regre          (Outcome="Disability", Age="Age", Sex="Sex", Cova=c("Education1",
                           "Education2", "Depression", "Diabetes", "Hypertension",
                           "RespiratorySymptoms", "MusculosqueletalSymptoms"), Step=FALSE,
                           CubAge=FALSE, InteractAS=FALSE, InteractWithAS=NULL, design=mod)
Sa=BurdenIndiv(Regression1, data=dat, SW="factor_f1")
write.table(Sa, dec=",")
```

Supplementary material chapter 4

Contents

Main functions for R S4.1

Table S4.1. Description of missing data. Chilean National Health Survey 2009-2020 (n=5.412).

Variable	N of missing data	%
Sample weight	75	1.4
Dsiability	119	2.2
Age	0	0.0
Sex	0	0.0
Education level	118	2.2
Income	355	6.6
Marital status	112	2.1
Hypertension	445	8.2
Diabetes	733	13.5
Chronic respiratory symptoms	119	2.2
Chronic musculoskeletal pain	294	5.4
Depresive episode	0	0.0
Total	1,138	21.0

Table S4.2. Effective coverage and other parameters for depressive episode estimated using the individual level approach applied on regression 2, **scenario 2** (**normative need: first episode in last 2 years/ utilization: 'treatment ever'**). Data from the Chilean National Health Survey 2009-2020 (n=4,447).

Prevalence 3.6 % [2.6 - 4.5]

Utilization 74.6% [63.2 - 86.1]

	Without constraining HG		Negative HGs are assumed equal to zero		Distribution of HG shifted up positive numbers	
	Individual level approach		Individual level approach		Individual level approach	
	model 2		model 2		model 2	
	mean	**UI**	**mean**	**UI**	**mean**	**UI**
HG_{G2}	-0.6	[-25.3 to 26.4]	5.6	[0.0 to 26.4]	31.9	[7.1 to 58.8]
HG_{G2}max	20.0	-	20.0	-	52.4	-
Relative benefit (%)	9.8	[-24.1 - 46.4]	52.2	[24.4 - 91.4]	284.3	[174.5 - 451]
Quality (%)	5.8	[-14.4 - 26.2]	30.8	[15.4 - 48.7]	64.6	[43.5 - 90.8]
relative - effective coverage (%)	7.6	[-18.6 – 35.0]	40.2	[19.3 - 67.4]	218.8	[137.9 - 330.2]
absolute - effective coverage (%)	0.1	[-0.3 - 0.5]	0.5	[0.3 - 0.8]	3.0	[2.1 - 3.8]
Effective coverage (%)	4.3	[-10.7 - 19.4]	23.0	[11.6 - 36.1]	48.1	[32.8 - 65.7]
$HG'_{G1} + HG'_{G2}$	61913	[50094 - 73667]	61846	[49864 - 73805]	161246	[129939 - 192724]
HG_{G2}	2648	[-6557 - 11853]	14157	[7218 - 21095]	76712	[55658 - 97766]
$RR_{HG\text{-}average}$	1.10	[0.76 - 1.46]	1.52	[1.24 - 1.91]	3.84	[2.75 - 5.51]
$RR_{HG\text{-}max}$	1.60	[1.37 - 1.89]	1.61	[1.37 - 1.9]	1.23	[1.14 - 1.35]

UI: uncertainty intervals

Table S4.3. Effective coverage and other parameters for depressive episode estimated using the individual level approach applied on regression 2, **scenario 3** (**normative need: all cases / utilization: 'medicaments last 2 weeks'**). Data from the Chilean National Health Survey 2009-2020 (n=4,447).

Prevalence 18.1 % [16.1 - 20.1]

Utilization 19.5% [14.8 - 24.2]

	Without constraining HG		Negative HGs are assumed equal to zero		Distribution of HG shifted up positive numbers	
	Individual level approach		Individual level approach		Individual level approach	
	model 2		model 2		model 2	
	mean	**UI**	**mean**	**UI**	**mean**	**UI**
HG_{G2}	-2.9	[-12.8 to 6.6]	0.8	[0 to 6.6]	11.9	[2.1 to 21.4]
$HG_{G2}max$	4.8	-	4.8	-	19.7	-
Relative benefit (%)	-23.7	[-39.4 - -10.2]	10.1	[3.7 - 17.2]	112.3	[77.5 - 154]
Quality (%)	-52.1	[-86.1 - -22.3]	22.4	[8.3 - 38.3]	62.5	[42.8 - 86.2]
relative - effective coverage (%)	-4.5	[-7.1 - -1.9]	1.9	[0.7 - 3.2]	21.2	[15.3 - 27.4]
absolute - effective coverage (%)	-0.3	[-0.4 - -0.1]	0.1	[0.0 - 0.2]	1.3	[1.0 - 1.7]
Effective coverage (%)	-9.8	[-15.5 - -4.3]	4.2	[1.6 - 6.9]	11.7	[8.5 - 15.2]
$HG'_{G1} + HG'_{G2}$	93996	[85012 - 102741]	93931	[85037 - 102766]	371377	[335559 - 406700]
HG_{G2}	-9201	[-14393 - -4010]	3961	[1529 - 6393]	43257	[31886 - 54628]
$RR_{HG\text{-}average}$	0.76	[0.61 - 0.9]	1.10	[1.04 - 1.17]	2.12	[1.78 - 2.54]
$RR_{HG\text{-}max}$	3.21	[2.61 - 4.02]	3.20	[2.61 - 4.02]	1.56	[1.41 - 1.76]

UI: uncertainty intervals

Table S4.4. Effective coverage and other parameters for depressive episode estimated using the individual level approach applied on regression 2, **scenario 3** (**normative need: severe cases / utilization: 'medicaments last 2 weeks'**). Data from the Chilean National Health Survey 2009-2020 (n=4,447).

Prevalence 4.0 % [2.9 - 5.1]

Utilization 22.1% [10.8 - 33.3]

	Without constraining HG		Negative HGs are assumed equal to zero		Distribution of HG shifted up positive numbers	
	Individual level approach		Individual level approach		Individual level approach	
	model 2		model 2		model 2	
	mean	**UI**	**mean**	**UI**	**mean**	**UI**
HG_{G2}	-1.6	[-32.2 to 30.3]	6.9	[0 to 30.3]	51.5	[20.8 to 83.4]
$HG_{G2}max$	25.5	-	25.5	-	78.6	-
Relative benefit (%)	-48.8	[-154.9 - 38.3]	60.7	[14.3 - 136.8]	667.9	[280.4 - 1425.4]
Quality (%)	-14.7	[-45.5 - 11.5]	18.5	[4.4 - 39.8]	66.7	[30.0 - 129.3]
relative - effective coverage (%)	-6.9	[-19.8 - 5.6]	8.7	[2.3 - 16.1]	96.4	[46.7 - 159.5]
absolute - effective coverage (%)	-0.1	[-0.3 - 0.1]	0.1	[0.0 - 0.2]	1.5	[0.8 - 2.2]
Effective coverage (%)	-3.1	[-8.7 - 2.5]	3.9	[1.0 - 7.1]	14.1	[7.0 - 22.4]
$HG'_{G1} + HG'_{G2}$	89011	[69854 - 108433]	89148	[69715 - 108383]	272419	[212577 - 333594]
HG_{G2}	-2737	[-7563 - 2089]	3441	[950 - 5931]	37847	[19333 - 56361]
$RR_{HG\text{-}average}$	0.51	[-0.55 - 1.38]	1.61	[1.14 - 2.37]	7.68	[3.80 - 15.25]
$RR_{HG\text{-}max}$	1.32	[1.14 - 1.65]	1.32	[1.14 - 1.63]	1.10	[1.05 - 1.21]

UI: uncertainty intervals

Main functions for R, used to estimate the effective coverage

#Funtion to simulate numbers with Beta distribution

```
# Input: mu (mean)
# Input: Va (variance)
# Input: n (number of simulated values)

DistrBETA=function(mu,Va,n) {
                                    alpha=(((1-mu)/Va)- (1/mu))*mu^2; alpha=abs(alpha)
                                    beta=alpha*((1/mu)-1); beta=abs(beta)
                                    mapply(rbeta, n=n, alpha, beta)}
```

```
#Function to calculate Effective Coverage through population – average level approach

library(survey)

# Input: G1 (name of the variable for G1)
# Input: G2 (name of the variable for G2)
# Input: Regression (the regression model as a text (as.character))
# Input: data (the database)
# Input: Sex (name of the variable for Sex)
# Input: Age (name of the variable for Age, non-centred and in x/10 scale)
# Input: Severity (name of the variable for Severity)
# Input: We (name of the variable for sample weights)


ECPop= function (G1, G2, Regression, data, Sex, Age, Severity, We) {

                #Editing inputs
                Vari=paste(as.character(Regression), collapse=" ")
                Vari=unlist(strsplit(Vari, " "))
                Vari=Vari[-1* c(grep("\\~", Vari), grep("\\+", Vari), grep("\\:", Vari), grep("\\(", Vari))]

                #Normative Need
                Disease=ifelse (data[,G1]==1 | data[,G2]==1,1,0)
                data=cbind(data, Disease)
                mod<-svydesign(~1, weights=as.formula(paste0("~",We)), data=data)
                NN=Preval("Disease",Sex=Sex, Age=Age, design=mod)
                nn=DistrBETA(NN$Prev[30], ((NN$Prev[3]- NN$Prev[30])/1.96)^2,10000)

                #Utilization
                data12=data[data$Disease==1,]
                mod12=svydesign(~1, weights=as.formula(paste0("~",We)), data=data12)
                Ut=Preval(G2,Sex, Age, mod12)
                ut=DistrBETA(Ut$Prev[30], ((Ut$Prev[3]- Ut$Prev[30])/1.96)^2,10000)

                #Model
                model=svyglm(as.formula(Regression), mod)

                # Datasets and counterfactuals
                # D0: without Normative Need
                da0=data[data[,G1]==0 & data[,G2]==0,]
                mod0=svydesign(~1, weights= as.formula(paste0("~",We)), data=da0)
                da0=data.frame(t(sapply(Vari, PopMeans, mod0)))


                # D1: with Normative Need, without Utilization
                da1=data[data[,G1]==1 & data[,G2]==0,]
                mod1=svydesign(~1, weights=as.formula(paste0("~",We)), data=da1)
                da1=data.frame(t(sapply(Vari, PopMeans, mod1)))
                da1prima=da1;                 da1prima[,G1]=0;
                da1prima[,Severity]=da0[,Severity]

                # D2 with Normative Need, with Utilization
                da2=data[data[,G1]==0 & data[,G2]==1,]
                mod2=svydesign(~1, weights= as.formula(paste0("~",We)), data=da2)
                da2=data.frame(t(sapply(Vari, PopMeans, mod2)))
                da2prima1=da2;                da2prima1[,G2]=0;    da2prima1[,Severity]=da0[,Severity]
```

```
da2prima2=da2;              da2prima2[,G2]=0;          da2prima2[,G1]=1

# Predicted values

D0= predict(model, da0);              D0=mapply(rnorm,10000, coef(D0), SE(D0))
D1= predict(model, da1);              D1=mapply(rnorm,10000, coef(D1), SE(D1))
D1p= predict(model, da1prima);        D1p=mapply(rnorm,10000, coef(D1p), SE(D1p))
D2= predict(model, da2);              D2=mapply(rnorm,10000, coef(D2), SE(D2))
D2p1= predict(model, da2prima1);      D2p1=mapply(rnorm,10000, coef(D2p1),
                                      SE(D2p1))
D2p2= predict(model, da2prima2);      D2p2=mapply(rnorm,10000, coef(D2p2),
                                      SE(D2p2))

# Efective Coverage
DA1=D1 - D1p
DA2=D2p2 - D2p1
HG=D2p2 - D2
RB=HG/DA2
ECr=(HG*ut) / ((DA1*(1-ut)) + (DA2*ut))
ECa=(HG*ut*nn) / ( (nn*((D1*(1-ut)) + (D2p2*ut))) + (D0*(1-nn)) )

# Output
Sa=list(D0=D0, D1=D1, D1p=D1p, DA1=DA1, D2=D2, D2p1=D2p1, D2p2=D2p2,
DA2=DA2, HG=HG, RB=RB, ECr=ECr, ECa=ECa)
for (k in names(Sa)) {Sa[[k]] = c(mean(Sa[[k]]), quantile(Sa[[k]], c(0.025, 0.975))) }
Output=matrix(unlist(Sa), ncol=3, byrow=TRUE)
rownames(Output)=names(Sa)
colnames(Output)=c("mean", "p2.5", "p97.5")
return(Output)}
```

```
#Function to calculate Effective Coverage through individual level approach

library(survey)

# Input: G1 (name of the variable for G1)
# Input: G2 (name of the variable for G2)
# Input: Regression (the regression model as a text (as.character))
# Input: data (the database)
# Input: Sex (name of the variable for Sex)
# Input: Age (name of the variable for Age, non-centred and in x/10 scale)
# Input: Severity (name of the variable for Severity)
# Input: We (name of the variable for sample weights)
# Input: Threshold (quantile that will be used to choose the maximum health gain)


ECIndiv= function (G1, G2, Regression, data, Sex, Age, Severity, We, Threshold) {

                #Editing inputs
                modI=svydesign(~1, weights=as.formula(paste0("~", We)), data=data)
                 assign("modI", modI, envir = globalenv())
                Vari=paste(as.character(Regression), collapse=" ")
                Vari=unlist(strsplit(Vari, " "))
                Vari=Vari[-1* c(grep("\\~", Vari), grep("\\+", Vari), grep("\\:", Vari), grep("\\(", Vari))]
                Outcome=Vari[30]
                Sev0=svyby(     as.formula(paste0("~", Severity)),
                                as.formula(paste0("~", G1, "==0 & ", G2, "==0")), modI, svymean)
                Sev0=Sev0[2,2]


                #Model
                model=svyglm(as.formula(Regression), modI)

                #Databases and counterfactuals
                da0=data[data[,G1]==0 & data[,G2]==0,]

                da1=data[data[,G1]==1 & data[,G2]==0,]
                da1prima1=da1;          da1prima1[,G1]=0;       da1prima1[,Severity]=Sev0
                da1prima2=da1;          da1prima2[,G1]=0;       da1prima2[,G2]=1

                da2=data[data[,G1]==0 & data[,G2]==1,]
                da2prima1=da2;          da2prima1[,G2]=0;       da2prima1[,Severity]=Sev0
                da2prima2=da2;          da2prima2[,G2]=0;       da2prima2[,G1]=1

                mod0=svydesign(~1, weights=as.formula(paste0("~", We)), data=da0)
                mod1=svydesign(~1, weights=as.formula(paste0("~", We)), data=da1)
                mod2=svydesign(~1, weights=as.formula(paste0("~", We)), data=da2)

                # Predicted values
                D0= predict(model, data.frame(da0))
                D1= predict(model, data.frame(da1))
                D1p1= predict(model, data.frame(da1prima1))
                D1p2= predict(model, data.frame(da1prima2))
                D2= predict(model, data.frame(da2))
                D2p1= predict(model, data.frame(da2prima1))
                D2p2= predict(model, data.frame(da2prima2))

                # Other parameters
```

```
DA1 = D1 - D1p1
DA2 = D2p2 - D2p1
RD1 = D1p2 - D1p1
RD2 = D2 - D2p1
HG2= D2p2 - D2
HG1= D1 - D1p2
HGmax=svyquantile(~HG2, mod2, Threshold)
HG2prima=ifelse(HG2<rep(HGmax, length(HG2)), HGmax, HG2)
HG1prima=ifelse(HG1<rep(HGmax, length(HG1)), HGmax, HG1)

# Sum of disability Parameters
D0xq=svytotal(~D0, mod0)
D1xq=svytotal(~D1, mod1)
DA1xq=svytotal(~DA1, mod1)
D2prima2xq= svytotal(~D2p2, mod2)
DA2xq=svytotal(~DA2, mod2)
HG2xq= svytotal(~HG2, mod2)
HG1primaxq= svytotal(~HG1prima, mod1)
HG2primaxq= svytotal(~HG2prima, mod2)


D0x=mapply(rnorm,10000, coef(D0xq), SE(D0xq))
D1x=mapply(rnorm,10000, coef(D1xq), SE(D1xq))
DA1x=mapply(rnorm,10000, coef(DA1xq), SE(DA1xq))
HG1primax= mapply(rnorm,10000, coef(HG1primaxq), SE(HG1primaxq))
HG2primax= mapply(rnorm,10000, coef(HG2primaxq), SE(HG2primaxq))
D2prima2x=mapply(rnorm,10000, coef(D2prima2xq), SE(D2prima2xq))
DA2x=mapply(rnorm,10000, coef(DA2xq), SE(DA2xq))
HG2x=mapply(rnorm,10000, coef(HG2xq), SE(HG2xq))


# Efective Coverage
HGmaxtot=HG1primax+HG2primax
HGrmax=svymean(~I(HG2/c(HGmax)), mod2)
HGr= svymean(~I(HG2/DA2), mod2, na.rm=TRUE)
Qual=HG2x/HG2primax
RB=HG2x/DA2x
ECr= HG2x / (DA1x + DA2x)
ECa= HG2x/ (D0x + D1x + D2prima2x)
ECr2=HG2x/ (HG1primax + HG2primax)


# Table Results 1
Sa=list(D0=D0, D1=D1, D1p1=D1p1, D1p2=D1p2, DA1=DA1, RD1=RD1, D2=D2,
D2p1=D2p1, D2p2=D2p2, DA2=DA2, RD2=RD2, HG1=HG1, HG2=HG2, ECr=ECr,
ECa=ECa, ECr2=ECr2, RB=RB, Qual=Qual)
for (k in names(Sa)) {Sa[[k]] = c(mean(Sa[[k]]), quantile(Sa[[k]], c(0.025, 0.975))) }

HGrmax=         c(coef(HGrmax),  coef(HGrmax)-1.96*SE(HGrmax),
                coef(HGrmax)+1.96*SE(HGrmax))
HGr=            c(coef(HGr), coef(HGr)-1.96*SE(HGr), coef(HGr)+1.96*SE(HGr))
HGmax=c(HGmax, NA,NA)

Sa=list(  D0=Sa["D0"], D1=Sa["D1"], D1p1=Sa["D1p1"],
          D1p2=Sa["D1p2"], DA1=Sa["DA1"], RD1=Sa["RD1"], D2=Sa["D2"],
          D2p1=Sa["D2p1"], D2p2=Sa["D2p2"], DA2=Sa["DA2"], RD2=Sa["RD2"],
          HG1=Sa["HG1"], HG2=Sa["HG2"], HGmax=HGmax, RB= Sa["RB"],
```

```
            Qual=Sa["Qual"], ECr=Sa["ECr"], ECa=Sa["ECa"], ECr2=Sa["ECr2"], HGr=HGr,
            HGrmax=HGrmax)

        TableR=matrix(unlist(Sa), ncol=3, byrow=TRUE)
        rownames(TableR)=names(Sa)
        colnames(TableR)=c("mean", "p2.5", "p97.5")


        # Table Results 2
        Sa=list(  D0=D0xq, D1=D1xq, DA1=DA1xq, D2p2=D2prima2xq, DA2=DA2xq,
            HG2=HG2xq, HG1p=HG1primaxq, HG2p=HG2primaxq)

        for (k in names(Sa)) {      Sa[[k]] = c(coef(Sa[[k]]), coef(Sa[[k]])-1.96*SE(Sa[[k]]),
                                    coef(Sa[[k]])+1.96*SE(Sa[[k]])) }
         HGmaxtot=      c(mean(HGmaxtot), quantile(HGmaxtot,c (0.025, 0.975)))

        Sa=list(  D0=Sa["D0"], D1=Sa["D1"], DA1=Sa["DA1"], D2p2=Sa["D2p2"],
            DA2=Sa["DA2"], HG2=Sa["HG2"], HG1p=Sa["HG1p"], HG2p=Sa["HG2p"],
            HGmaxtot=HGmaxtot)

        TableR2=matrix(unlist(Sa), ncol=3, byrow=TRUE)
        rownames(TableR2)=names(Sa)
        colnames(TableR2)=c("mean", "p2.5", "p97.5")


        # Distributions
        L=length(Vari)+1
        da0=da0[,c(Vari,We)];     da0=cbind(da0,D0)
        names(da0)[c(L+1,L+2)]=c("D0", "D0se")

        da1=da1[,c(Vari,We)]; da1=cbind(da1,D1, D1p1, D1p2, DA1, RD1, HG1, HG1prima)

        names(da1)[(L+1):ncol(da1)]=      c("D1", "D1se", "D1p1", "D1p1se", "D1p2",
                                          "D1p2se", "DA1", "DA1se","RD1", "RD1se",
                                          "HG1", "HG1se", "HG1prima")

        da2=da2[,c(Vari,We)];     da2=cbind(da2,D2, D2p1, D2p2, DA2, RD2, HG2,
HG2prima)
        names(da2)[(L+1):ncol(da2)]=      c("D2", "D2se", "D2p1", "D2p1se", "D2p2",
                                          "D2p2se", "DA2", "DA2se", "RD2", "RD2se",
                                          "HG2", "HG2se", "HG2prima")

        Distributions=list(da0=da0, da1=da1, da2=da2)

        # Output
        Output=list(TableR=TableR, TableR2=TableR2, Distributions=Distributions)
        return(Output)}
```

Supplementary material chapter 5

Contents

Table S5.1. Description of Health State Description Questionnaire. China. SAGE study. wave 1. (n=14.248).

		Category																
Item	Description	0	1	2	3	4	**n**	**mean**	**sd**	**median**	**trimmed**	**mad**	**min**	**max**	**range**	**skew**	**kurtosis**	**se**
q2000	Health today	637	4634	6270	2437	270	14248	1.8	0.8	2	1.78	1.48	0	4	4	0.14	-0.20	0.01
q2001	Difficulty 30 days	9135	3428	1262	355	68	14248	0.5	0.8	0	0.35	0.00	0	4	4	1.63	2.36	0.01
q2002	Moving around	11538	1926	588	170	26	14248	0.3	0.6	0	0.11	0.00	0	4	4	2.67	7.66	0.01
q2003	Vigorous activities	5505	3849	2807	1537	549	14247	1.1	1.2	1	1.00	1.48	0	4	4	0.73	-0.43	0.01
q2004	Self-care	13087	823	232	81	25	14248	0.1	0.4	0	0.00	0.00	0	4	4	4.71	25.97	0.00
q2005	General appearance	13230	737	197	61	23	14248	0.1	0.4	0	0.00	0.00	0	4	4	5.11	31.14	0.00
q2006	Staying by yourself	12907	853	297	116	75	14248	0.1	0.5	0	0.00	0.00	0	4	4	4.45	22.38	0.00
q2007	Body pain	7871	4389	1614	350	24	14248	0.6	0.8	0	0.48	0.00	0	4	4	1.17	0.81	0.01
q2008	Body discomfort	7620	4669	1641	294	24	14248	0.6	0.8	0	0.50	0.00	0	4	4	1.10	0.71	0.01
q2009	Difficulty because of pain	8537	4049	1366	268	28	14248	0.5	0.8	0	0.40	0.00	0	4	4	1.35	1.42	0.01
q2010	Concentrating and remembering	8172	4308	1434	315	18	14247	0.6	0.8	0	0.44	0.00	0	4	4	1.25	1.05	0.01
q2011	Learning	6767	4631	2146	640	64	14248	0.8	0.9	1	0.66	1.48	0	4	4	0.98	0.29	0.01
q2012	Participation	12927	1017	237	59	8	14248	0.1	0.4	0	0.00	0.00	0	4	4	4.13	19.98	0.00
q2013	Dealing with conflicts	12731	1159	287	64	7	14248	0.1	0.4	0	0.01	0.00	0	4	4	3.75	16.08	0.00
q2014	New and maintain friends	12971	985	228	57	7	14248	0.1	0.4	0	0.00	0.00	0	4	4	4.20	20.62	0.00
q2015	Dealing with strangers	11048	1688	842	535	134	14247	0.4	0.8	0	0.16	0.00	0	4	4	2.33	4.87	0.01
q2016	Sleep	8954	3594	1361	308	31	14248	0.5	0.8	0	0.36	0.00	0	4	4	1.47	1.67	0.01
q2017	Rested and refreshed	9231	3710	1092	203	12	14248	0.5	0.7	0	0.32	0.00	0	4	4	1.51	1.88	0.01
q2018	Felling sad	11705	1985	443	107	8	14248	0.2	0.5	0	0.10	0.00	0	4	4	2.70	7.99	0.00
q2019	Anxiety	11657	2013	470	99	5	14244	0.2	0.5	0	0.10	0.00	0	4	4	2.61	7.27	0.00
q2023	Vision across the road	9282	3248	1295	354	69	14248	0.5	0.8	0	0.34	0.00	0	4	4	1.65	2.39	0.01
q2024	Vision at arm's length	6809	4689	2245	425	80	14248	0.8	0.9	1	0.64	1.48	0	4	4	0.98	0.44	0.01

sd: standard deviation / min: minimum / max: maximum / se: standard error

Table S5.2. Description of Health State Description Questionnaire. Ghana. SAGE study. wave 1.

		Category															
Item	Description	0	1	2	3	4	**n**	**mean**	**sd**	**median**	**trimmed**	**mad**	**min**	**max**	**range**	**skew**	**kurtosis**	**se**
q2000	Health today	345	1988	1992	672	106	5103	1.6	0.9	2	1.62	1.48	0	4	4	0.31	-0.08	0.01
q2001	Difficulty 30 days	1210	1221	1928	617	117	5093	1.5	1.1	2	1.41	1.48	0	4	4	0.12	-0.73	0.01
q2002	Moving around	2914	1082	787	274	46	5103	0.7	1.0	0	0.56	0.00	0	4	4	1.18	0.47	0.01
q2003	Vigorous activities	1415	789	1227	977	696	5104	1.8	1.4	2	1.69	1.48	0	4	4	0.13	-1.25	0.02
q2004	Self-care	3924	758	311	74	37	5104	0.3	0.7	0	0.16	0.00	0	4	4	2.43	6.26	0.01
q2005	General appearance	3914	709	360	85	35	5103	0.4	0.7	0	0.17	0.00	0	4	4	2.32	5.40	0.01
q2006	Staying by yourself	3614	707	464	171	145	5101	0.5	1.0	0	0.31	0.00	0	4	4	1.95	3.15	0.01
q2007	Body pain	1298	1715	1389	683	17	5102	1.3	1.0	1	1.24	1.48	0	4	4	0.26	-0.93	0.01
q2008	Body discomfort	1412	1642	1404	627	16	5101	1.3	1.0	1	1.19	1.48	0	4	4	0.28	-0.94	0.01
q2009	Difficulty because of pain	1599	1719	1239	462	19	5038	1.1	1.0	1	1.03	1.48	0	4	4	0.44	-0.73	0.01
q2010	Concentrating and remembering	2144	1524	1103	321	10	5102	0.9	0.9	1	0.83	1.48	0	4	4	0.64	-0.62	0.01
q2011	Learning	2366	1376	1049	269	41	5101	0.9	1.0	1	0.75	1.48	0	4	4	0.82	-0.20	0.01
q2012	Participation	3169	896	631	304	102	5102	0.7	1.0	0	0.48	0.00	0	4	4	1.43	1.13	0.01
q2013	Dealing with conflicts	3445	796	505	271	85	5102	0.6	1.0	0	0.37	0.00	0	4	4	1.68	1.97	0.01
q2014	New and maintain friends	3379	845	585	218	71	5098	0.6	0.9	0	0.39	0.00	0	4	4	1.62	1.87	0.01
q2015	Dealing with strangers	3425	758	619	234	59	5095	0.6	0.9	0	0.38	0.00	0	4	4	1.58	1.62	0.01
q2016	Sleep	2340	1414	1028	309	13	5104	0.9	1.0	1	0.76	1.48	0	4	4	0.75	-0.48	0.01
q2017	Rested and refreshed	2339	1320	1137	298	10	5104	0.9	1.0	1	0.78	1.48	0	4	4	0.68	-0.67	0.01
q2018	Felling sad	2579	1534	805	178	5	5101	0.7	0.9	0	0.61	0.00	0	4	4	0.93	-0.04	0.01
q2019	Anxiety	2226	1431	1067	356	10	5090	0.9	1.0	1	0.81	1.48	0	4	4	0.68	-0.62	0.01
q2023	Vision across the road	1987	1363	1273	425	50	5098	1.1	1.0	1	0.94	1.48	0	4	4	0.56	-0.67	0.01
q2024	Vision at arm's length	2090	1480	1085	399	40	5094	1.0	1.0	1	0.86	1.48	0	4	4	0.69	-0.48	0.01

sd: standard deviation / min: minimum / max: maximum / se: standard error

Table S5.3. Description of Health State Description Questionnaire. India. SAGE study. wave 1.

		Category				
Item	Description	0	1	2	3	4
q2000	Health today	550	3824	5080	1626	147
q2001	Difficulty 30 days	3200	2884	3538	1355	251
q2002	Moving around	6308	2479	1291	998	152
q2003	Vigorous activities	3663	2486	1840	2110	1085
q2004	Self-care	9346	1185	415	222	60
q2005	General appearance	9562	1035	388	189	52
q2006	Staying by yourself	7369	2070	902	630	257
q2007	Body pain	3997	3518	2091	1503	118
q2008	Body discomfort	4132	3373	2201	1378	143
q2009	Difficulty because of pain	4241	3256	2322	1244	131
q2010	Concentrating and remembering	5002	3350	1865	889	117
q2011	Learning	4825	2901	1836	1244	420
q2012	Participation	6669	2705	1177	430	244
q2013	Dealing with conflicts	6229	2473	1605	611	307
q2014	New and maintain friends	7004	2236	1358	484	142
q2015	Dealing with strangers	6110	2493	1415	900	304
q2016	Sleep	5951	2518	1629	1047	79
q2017	Rested and refreshed	4974	3170	1849	1125	107
q2018	Felling sad	5486	3264	1597	814	63
q2019	Anxiety	4540	3283	1857	1377	120
q2023	Vision across the road	5421	2596	1893	1151	159
q2024	Vision at arm's length	5432	2825	1863	975	125

n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
11227	1.7	0.8	2	1.71	1.48	0	4	4	0.13	-0.12	0.01
11228	1.3	1.1	1	1.27	1.48	0	4	4	0.29	-0.79	0.01
11228	0.8	1.0	0	0.57	0.00	0	4	4	1.21	0.40	0.01
11184	1.5	1.4	1	1.39	1.48	0	4	4	0.40	-1.16	0.01
11228	0.3	0.7	0	0.08	0.00	0	4	4	3.02	9.56	0.01
11226	0.2	0.6	0	0.06	0.00	0	4	4	3.24	11.16	0.01
11228	0.6	1.0	0	0.38	0.00	0	4	4	1.72	2.17	0.01
11227	1.1	1.1	1	1.02	1.48	0	4	4	0.59	-0.72	0.01
11227	1.1	1.1	1	1.00	1.48	0	4	4	0.61	-0.66	0.01
11194	1.1	1.1	1	0.97	1.48	0	4	4	0.62	-0.64	0.01
11223	0.9	1.0	1	0.76	1.48	0	4	4	0.89	-0.08	0.01
11226	1.1	1.2	1	0.91	1.48	0	4	4	0.84	-0.34	0.01
11225	0.7	1.0	0	0.46	0.00	0	4	4	1.58	2.06	0.01
11225	0.8	1.1	0	0.59	0.00	0	4	4	1.29	0.87	0.01
11224	0.6	0.9	0	0.44	0.00	0	4	4	1.49	1.52	0.01
11222	0.8	1.1	0	0.62	0.00	0	4	4	1.21	0.48	0.01
11224	0.8	1.0	0	0.64	0.00	0	4	4	1.01	-0.14	0.01
11225	1.0	1.0	1	0.80	1.48	0	4	4	0.83	-0.36	0.01
11224	0.8	1.0	1	0.66	1.48	0	4	4	1.00	0.10	0.01
11177	1.0	1.1	1	0.91	1.48	0	4	4	0.72	-0.58	0.01
11220	0.9	1.1	1	0.77	1.48	0	4	4	0.87	-0.36	0.01
11220	0.9	1.0	1	0.72	1.48	0	4	4	0.92	-0.18	0.01

sd: standard deviation / min: minimum / max: maximum / se: standard error

Table S5.4. Description of Health State Description Questionnaire. Mexico. SAGE study. wave 1.

		Category																
Item	Description	0	1	2	3	4	**n**	**mean**	**sd**	**median**	**trimmed**	**mad**	**min**	**max**	**range**	**skew**	**kurtosis**	**se**
q2000	Health today	96	983	1255	285	17	2636	1.7	0.7	2	1.63	1.48	0	4	4	0.15	-0.02	0.01
q2001	Difficulty 30 days	1305	601	477	156	97	2636	0.9	1.1	1	0.73	1.48	0	4	4	1.07	0.30	0.02
q2002	Moving around	1536	481	384	186	49	2636	0.8	1.1	0	0.57	0.00	0	4	4	1.22	0.48	0.02
q2003	Vigorous activities	1027	477	402	218	512	2636	1.5	1.5	1	1.39	1.48	0	4	4	0.53	-1.22	0.03
q2004	Self-care	2056	305	149	73	53	2636	0.4	0.9	0	0.16	0.00	0	4	4	2.49	5.85	0.02
q2005	General appearance	2192	247	121	46	30	2636	0.3	0.7	0	0.09	0.00	0	4	4	3.01	9.34	0.01
q2006	Staying by yourself	2066	253	115	64	138	2636	0.5	1.1	0	0.17	0.00	0	4	4	2.39	4.67	0.02
q2007	Body pain	1202	698	508	213	15	2636	0.9	1.0	1	0.78	1.48	0	4	4	0.78	-0.46	0.02
q2008	Body discomfort	1188	807	451	172	18	2636	0.9	1.0	1	0.74	1.48	0	4	4	0.88	-0.06	0.02
q2009	Difficulty because of pain	1372	687	414	147	16	2636	0.8	0.9	0	0.62	0.00	0	4	4	1.04	0.20	0.02
q2010	Concentrating and remembering	1329	756	470	75	6	2636	0.7	0.9	0	0.63	0.00	0	4	4	0.88	-0.13	0.02
q2011	Learning	1492	679	358	85	22	2636	0.7	0.9	0	0.51	0.00	0	4	4	1.28	1.09	0.02
q2012	Participation	1944	428	185	44	35	2636	0.4	0.8	0	0.20	0.00	0	4	4	2.29	5.43	0.02
q2013	Dealing with conflicts	1892	469	200	53	22	2636	0.4	0.8	0	0.23	0.00	0	4	4	2.07	4.26	0.02
q2014	New and maintain friends	2054	401	139	28	14	2636	0.3	0.7	0	0.15	0.00	0	4	4	2.52	7.05	0.01
q2015	Dealing with strangers	1927	470	182	40	17	2636	0.4	0.7	0	0.21	0.00	0	4	4	2.15	4.81	0.01
q2016	Sleep	1431	575	467	156	7	2636	0.8	1.0	0	0.62	0.00	0	4	4	0.96	-0.21	0.02
q2017	Rested and refreshed	1394	709	409	120	4	2636	0.7	0.9	0	0.59	0.00	0	4	4	1.00	0.04	0.02
q2018	Felling sad	1413	695	366	152	10	2636	0.7	0.9	0	0.58	0.00	0	4	4	1.10	0.30	0.02
q2019	Anxiety	1338	734	419	138	7	2636	0.8	0.9	0	0.63	0.00	0	4	4	0.97	0.01	0.02
q2023	Vision across the road	1497	637	363	116	23	2636	0.7	0.9	0	0.53	0.00	0	4	4	1.26	0.86	0.02
q2024	Vision at arm's length	1392	673	411	130	30	2636	0.8	1.0	0	0.61	0.00	0	4	4	1.13	0.54	0.02

sd: standard deviation / min: minimum / max: maximum / se: standard error

Table S5.5. Description of Health State Description Questionnaire. The Russian Federation SAGE study. wave 1.

		Category														
Item	Description	0	1	2	3	4	**n**	**mean**	**sd**	**median**	**trimmed**	**mad**	**min**	**max**	**range**	**skew**	**kurtosis**	**se**
q2000	Health today	37	652	2521	1039	84	4333	2.1	0.7	2	2.13	0.00	0	4	4	0.03	0.32	0.01
q2001	Difficulty 30 days	1068	1175	1440	548	92	4323	1.4	1.1	1	1.35	1.48	0	4	4	0.22	-0.75	0.02
q2002	Moving around	1911	1033	869	444	74	4331	1.0	1.1	1	0.87	1.48	0	4	4	0.74	-0.52	0.02
q2003	Vigorous activities	981	1199	1072	747	282	4281	1.6	1.2	1	1.50	1.48	0	4	4	0.31	-0.87	0.02
q2004	Self-care	3085	639	411	147	43	4325	0.5	0.9	0	0.28	0.00	0	4	4	1.88	2.93	0.01
q2005	General appearance	3447	484	304	68	24	4327	0.3	0.7	0	0.13	0.00	0	4	4	2.45	5.94	0.01
q2006	Staying by yourself	2865	674	335	132	43	4049	0.5	0.9	0	0.27	0.00	0	4	4	1.95	3.38	0.01
q2007	Body pain	1664	1201	951	452	58	4326	1.1	1.1	1	0.96	1.48	0	4	4	0.62	-0.63	0.02
q2008	Body discomfort	1600	1240	959	468	57	4324	1.1	1.1	1	0.99	1.48	0	4	4	0.59	-0.65	0.02
q2009	Difficulty because of pain	1787	1263	816	414	38	4318	1.0	1.0	1	0.86	1.48	0	4	4	0.73	-0.47	0.02
q2010	Concentrating and remembering	2065	1347	655	246	11	4324	0.8	0.9	1	0.67	1.48	0	4	4	0.94	0.02	0.01
q2011	Learning	2134	1216	604	313	44	4311	0.8	1.0	1	0.66	1.48	0	4	4	1.06	0.27	0.02
q2012	Participation	3152	710	285	110	56	4313	0.4	0.8	0	0.22	0.00	0	4	4	2.22	4.85	0.01
q2013	Dealing with conflicts	2814	1050	314	95	31	4304	0.5	0.8	0	0.31	0.00	0	4	4	1.82	3.41	0.01
q2014	New and maintain friends	3304	687	231	63	32	4317	0.3	0.7	0	0.17	0.00	0	4	4	2.48	6.73	0.01
q2015	Dealing with strangers	3014	883	295	82	30	4304	0.4	0.8	0	0.25	0.00	0	4	4	2.02	4.24	0.01
q2016	Sleep	1319	1500	943	475	48	4285	1.2	1.0	1	1.07	1.48	0	4	4	0.54	-0.57	0.02
q2017	Rested and refreshed	1160	1619	1066	391	40	4276	1.2	1.0	1	1.10	1.48	0	4	4	0.47	-0.46	0.01
q2018	Felling sad	2534	1004	528	164	22	4252	0.6	0.9	0	0.47	0.00	0	4	4	1.33	1.06	0.01
q2019	Anxiety	2126	1354	582	173	30	4265	0.7	0.9	1	0.61	1.48	0	4	4	1.12	0.73	0.01
q2023	Vision across the road	2300	1227	553	199	38	4317	0.7	0.9	0	0.56	0.00	0	4	4	1.23	0.91	0.01
q2024	Vision at arm's length	2176	1254	664	188	32	4314	0.8	0.9	0	0.63	0.00	0	4	4	1.06	0.46	0.01

sd: standard deviation / min: minimum / max: maximum / se: standard error

Table S5.6. Description of Health State Description Questionnaire. South Africa. SAGE study. wave 1.

		Category																
Item	Description	0	1	2	3	4	**n**	**mean**	**sd**	**median**	**trimmed**	**mad**	**min**	**max**	**range**	**skew**	**kurtosis**	**se**
q2000	Health today	250	1472	1762	580	65	4129	1.7	0.8	2	1.67	1.48	0	4	4	0.17	-0.14	0.01
q2001	Difficulty 30 days	1737	660	1325	337	46	4105	1.1	1.1	1	0.99	1.48	0	4	4	0.41	-1.01	0.02
q2002	Moving around	2944	481	499	182	26	4132	0.5	0.9	0	0.32	0.00	0	4	4	1.65	1.70	0.01
q2003	Vigorous activities	2299	535	658	402	222	4116	1.0	1.3	0	0.76	0.00	0	4	4	1.01	-0.25	0.02
q2004	Self-care	3566	271	230	53	12	4132	0.2	0.6	0	0.05	0.00	0	4	4	3.02	9.09	0.01
q2005	General appearance	3580	267	225	46	13	4131	0.2	0.6	0	0.04	0.00	0	4	4	3.09	9.70	0.01
q2006	Staying by yourself	3345	318	267	135	62	4127	0.4	0.9	0	0.13	0.00	0	4	4	2.50	5.62	0.01
q2007	Body pain	1605	1151	940	417	15	4128	1.1	1.0	1	0.94	1.48	0	4	4	0.54	-0.84	0.02
q2008	Body discomfort	1675	1121	949	367	14	4126	1.0	1.0	1	0.90	1.48	0	4	4	0.57	-0.80	0.02
q2009	Difficulty because of pain	1837	966	884	331	15	4033	0.9	1.0	1	0.81	1.48	0	4	4	0.68	-0.71	0.02
q2010	Concentrating and remembering	1979	916	960	259	9	4123	0.9	1.0	1	0.77	1.48	0	4	4	0.68	-0.75	0.02
q2011	Learning	2054	789	921	305	55	4124	0.9	1.1	1	0.77	1.48	0	4	4	0.81	-0.44	0.02
q2012	Participation	2944	564	436	152	33	4129	0.5	0.9	0	0.30	0.00	0	4	4	1.79	2.41	0.01
q2013	Dealing with conflicts	2852	600	506	152	17	4127	0.5	0.9	0	0.34	0.00	0	4	4	1.58	1.53	0.01
q2014	New and maintain friends	3028	526	397	147	27	4125	0.5	0.9	0	0.26	0.00	0	4	4	1.90	2.83	0.01
q2015	Dealing with strangers	2803	550	525	198	42	4118	0.6	1.0	0	0.38	0.00	0	4	4	1.55	1.44	0.01
q2016	Sleep	2340	760	674	332	17	4123	0.8	1.0	0	0.60	0.00	0	4	4	1.02	-0.20	0.02
q2017	Rested and refreshed	2339	798	690	276	15	4118	0.7	1.0	0	0.59	0.00	0	4	4	1.04	-0.10	0.02
q2018	Felling sad	2274	912	668	252	19	4125	0.7	1.0	0	0.60	0.00	0	4	4	1.05	0.05	0.02
q2019	Anxiety	2215	922	669	290	24	4120	0.8	1.0	0	0.63	0.00	0	4	4	1.02	-0.04	0.02
q2023	Vision across the road	2163	819	864	212	26	4084	0.8	1.0	0	0.68	0.00	0	4	4	0.89	-0.31	0.02
q2024	Vision at arm's length	2241	789	824	216	22	4092	0.8	1.0	0	0.64	0.00	0	4	4	0.94	-0.25	0.02

sd: standard deviation / min: minimum / max: maximum / se: standard error

Figures S5.1-6. Representation of polychoric correlation matrix of items from the Health State Description Questionnaire, in six countries. SAGE study. wave 1.

Figure S5.1. China

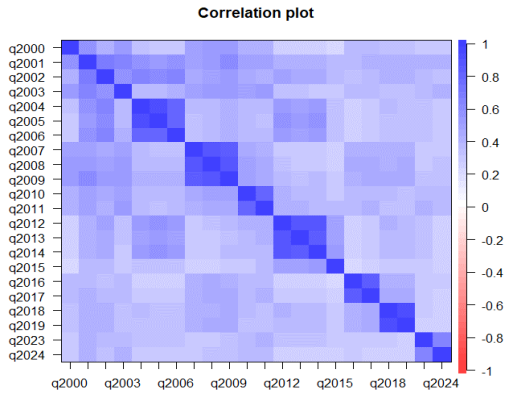


Figure S5.2. Ghana

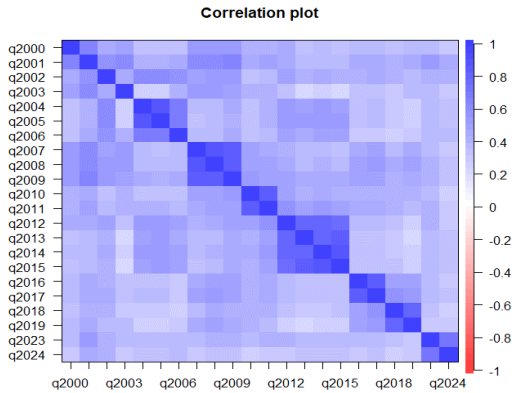


Figure S5.3. India

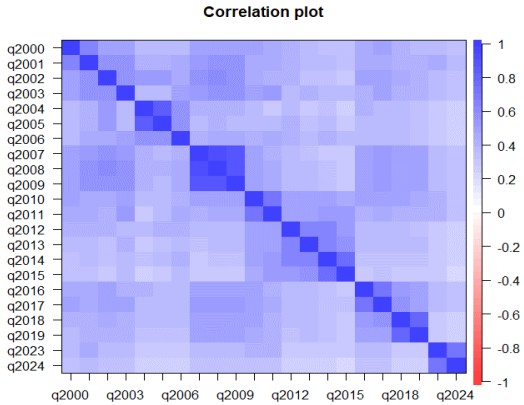


Figure S5.4. Mexico

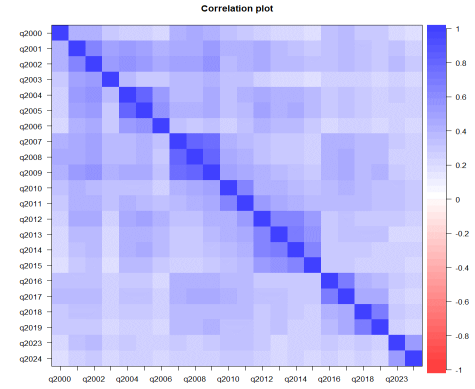


Figure S5.5. Russian Federation

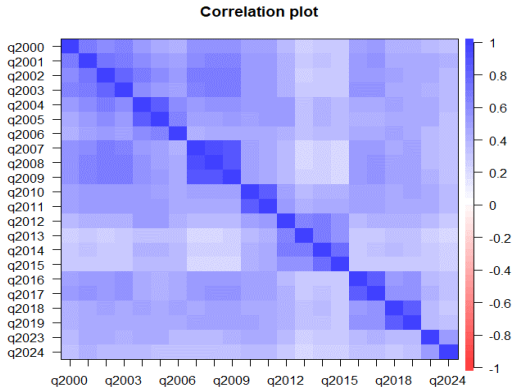
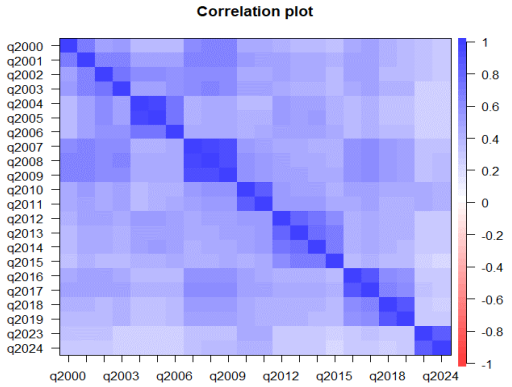


Figure S5.6. South Africa



Figures S5.7-12. Parallel Analysis to select the number of factors from the Health State Description Questionnaire, in six countries. SAGE study. wave 1.

Figure S5.7. China

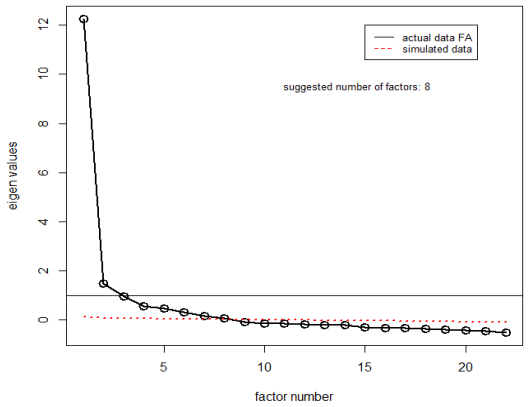


Figure S5.8. Ghana

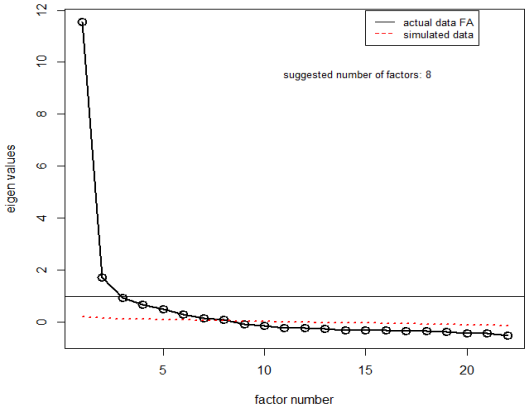


Figure S5.9. India

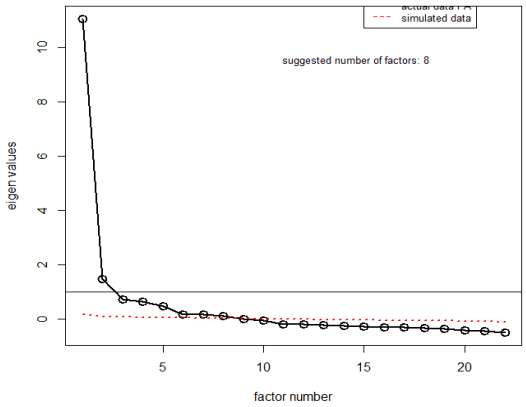


Figure S5.10. Mexico

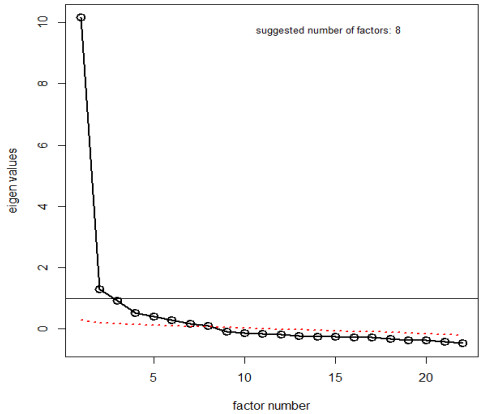


Figure S5.11. Russian Federation

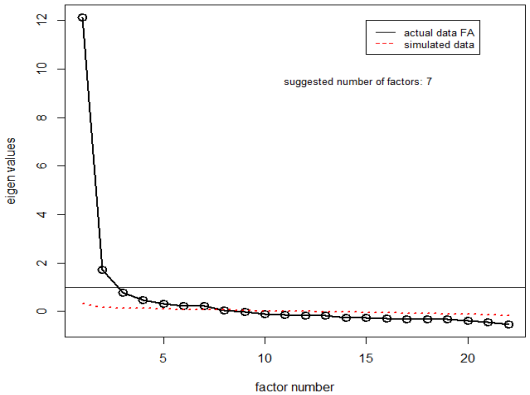


Figure S5.12. South Africa

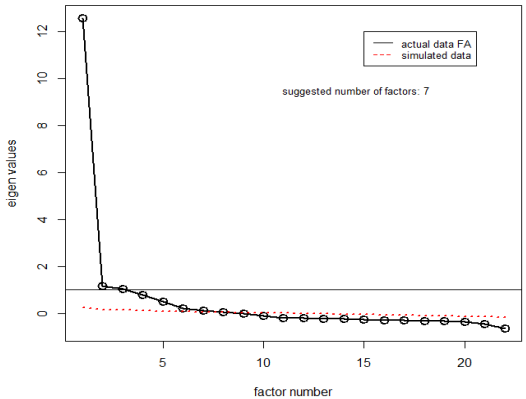


Table S5.7. Exploratory Factor Analysis of the Health State Description Questionnaire. China. SAGE study. wave 1.

	InterP	SelfC	Pain	Affect	SleepEng	Cognit	Vision	Gen-Mob	
	WLS5	**WLS1**	**WLS2**	**WLS3**	**WLS6**	**WLS7**	**WLS4**	**WLS8**	**Uniq**
q2000								0.30	0.48
q2001		*0.36*						*0.36*	0.19
q2002		*0.43*						*0.33*	0.17
q2003								0.46	0.29
q2004		**0.93**							0.04
q2005		**0.95**							0.03
q2006		**0.85**							0.10
q2007			**0.97**						0.10
q2008			**0.96**						0.06
q2009			**0.90**						0.07
q2010						**0.88**			0.17
q2011						**0.86**			0.16
q2012	**0.87**								0.05
q2013	**0.98**								0.07
q2014	**0.91**								0.05
q2015	**0.77**								0.42
q2016					**0.94**				0.14
q2017					**0.92**				0.12
q2018				**0.96**					0.07
q2019				**0.97**					0.07
q2023							**0.80**		0.29
q2024							**0.83**		0.33

n	7,115	RMSEA	0.10	
rotation	oblimin	lower	0.10	
method	wls	upper	0.10	
		confidence	0.9	

Factors: *Gen: general health/ Mob: mobility/ SelfC: Self-care/ Pain: Pain and Discomfort/ Cognit: cognition/ InterP: Interpersonal activities/ SleepEng: Sleep and energy/ Affect: affects/ Vision: vision*

Uniq: uniqueness.

Factor loadings lower than 0.2 are omitted.

Larger factor loadings are marked with bold.

Italic numbers show cross-loadings.

Table S5.8. Exploratory Factor Analysis of the Health State Description Questionnaire. Ghana. SAGE study. wave 1.

	InterP	Pain	SelfC-Mob	SleepEng	Affect	Vision	Cognit	Gen-Mob	
	WLS1	**WLS2**	**WLS8**	**WLS7**	**WLS3**	**WLS5**	**WLS4**	**WLS6**	**Uniq**
q2000								**0.52**	0.45
q2001								**0.55**	0.22
q2002			***0.50***					*0.32*	0.25
q2003								**0.64**	0.35
q2004			**0.96**						0.04
q2005			**0.93**						0.05
q2006			**0.68**						0.24
q2007		**1.00**							0.05
q2008		**0.93**							0.07
q2009		**0.83**							0.12
q2010							**0.84**		0.14
q2011							**0.78**		0.13
q2012	**0.82**								0.13
q2013	**0.90**								0.10
q2014	**0.90**								0.09
q2015	**0.93**								0.08
q2016				**0.94**					0.14
q2017				**0.91**					0.09
q2018					**0.90**				0.12
q2019					**0.88**				0.14
q2023						**0.82**			0.21
q2024						**0.88**			0.26

n	2,552	RMSEA	0.10	
rotation	oblimin	lower	0.10	
method	wls	upper	0.10	
		confidence	0.9	

Factors: *Gen: general health/ Mob: mobility/ SelfC: Self-care/ Pain: Pain and Discomfort/ Cognit: cognition/ InterP: Interpersonal activities/ SleepEng: Sleep and energy/ Affect: affects/ Vision: vision*

Uniq: uniqueness.

Factor loadings lower than 0.2 are omitted.

Larger factor loadings are marked with bold.

Italic numbers show cross-loadings.

Table S5.9. Exploratory Factor Analysis of the Health State Description Questionnaire. India. SAGE study. wave 1.

	Pain	InterP	SelfC-Mob	Affect	SleepEng	Vision	Cognit	Gen-Mob	
	WLS1	**WLS2**	**WLS6**	**WLS3**	**WLS8**	**WLS4**	**WLS7**	**WLS5**	**Uniq**
q2000								0.48	0.42
q2001								**0.60**	0.27
q2002			*0.35*					*0.32*	0.30
q2003							*0.29*	*0.37*	0.41
q2004			**0.95**						0.03
q2005			**0.93**						0.04
q2006			**0.57**						0.39
q2007	**1.01**								0.04
q2008	**0.96**								0.03
q2009	**0.88**								0.11
q2010							**0.70**		0.25
q2011							**0.82**		0.20
q2012		**0.70**							0.34
q2013		**0.78**							0.31
q2014		**0.91**							0.15
q2015		**0.85**							0.27
q2016					**0.85**				0.22
q2017					**0.86**				0.20
q2018				**0.89**					0.13
q2019				**0.90**					0.18
q2023						**0.85**			0.23
q2024						**0.87**			0.28

n	5,614	RMSEA	0.10
rotation	oblimin	lower	0.10
method	wls	upper	0.10
		confidence	0.9

Factors: *Gen: general health/ Mob: mobility/ SelfC: Self-care/ Pain: Pain and Discomfort/ Cognit: cognition/ InterP: Interpersonal activities/ SleepEng: Sleep and energy/ Affect: affects/ Vision: vision*

Uniq: uniqueness.

Factor loadings lower than 0.2 are omitted.

Larger factor loadings are marked with bold.

Italic numbers show cross-loadings.

Table S5.10. Exploratory Factor Analysis of the Health State Description Questionnaire. Mexico. SAGE study. wave 1.

	InterP	Pain	SelfC	Gen-Mob	Affect	SleepEng	Cognit	Vision	
	WLS5	**WLS2**	**WLS1**	**WLS8**	**WLS3**	**WLS7**	**WLS6**	**WLS4**	**Uniq**
q2000		*0.22*		*0.40*					0.62
q2001				**0.61**					0.34
q2002				**0.71**					0.17
q2003				**0.69**					0.47
q2004			**0.76**						0.13
q2005			**0.91**						0.01
q2006			**0.51**	*0.23*					0.34
q2007		**0.91**							0.18
q2008		**0.92**							0.13
q2009		**0.78**							0.17
q2010							**0.57**		0.40
q2011							**1.01**		0.00
q2012	**0.70**								0.20
q2013	**0.87**								0.20
q2014	**0.89**								0.18
q2015	**0.79**								0.32
q2016						**0.65**			0.43
q2017						**1.00**			0.00
q2018					**0.82**				0.23
q2019					**0.89**				0.19
q2023								**1.00**	0.00
q2024								0.47	0.63

n	1,318	RMSEA	0.06
rotation	oblimin	lower	0.06
method	wls	upper	0.07
		confidence	0.9

Factors: *Gen: general health/ Mob: mobility/ SelfC: Self-care/ Pain: Pain and Discomfort/ Cognit: cognition/ InterP: Interpersonal activities/ SleepEng: Sleep and energy/ Affect: affects/ Vision: vision*

Uniq: uniqueness.

Factor loadings lower than 0.2 are omitted.

Larger factor loadings are marked with bold.

Italic numbers show cross-loadings.

Table S5.11. Exploratory Factor Analysis of the Health State Description Questionnaire. Russian Federation. SAGE study. wave 1.

	Pain	InterP	SelfCGenMob	SleepEng	Cognit	Affect	Vision	
	WLS1	**WLS2**	**WLS6**	**WLS5**	**WLS4**	**WLS3**	**WLS7**	**Uniq**
q2000	*0.31*		*0.23*	*0.32*				0.36
q2001	*0.22*		*0.32*	*0.30*				0.26
q2002	*0.27*		**0.56**					0.17
q2003	*0.29*			*0.26*				0.25
q2004			**0.75**					0.09
q2005			**0.75**					0.09
q2006			**0.57**			*0.27*		0.34
q2007	**0.96**							0.08
q2008	**0.94**							0.06
q2009	**0.86**							0.10
q2010					**0.91**			0.14
q2011					**0.90**			0.14
q2012		**0.69**						0.24
q2013		**0.87**						0.27
q2014		**0.94**						0.12
q2015		**0.83**						0.24
q2016				**0.85**				0.18
q2017				**0.82**				0.16
q2018						**0.72**		0.14
q2019						**0.73**		0.14
q2023							**0.72**	0.44
q2024							**0.73**	0.51

n	2,169	RMSEA	0.11
rotation	oblimin	lower	0.11
method	wls	upper	0.12
		confidence	0.9

Factors: *Gen: general health/ Mob: mobility/ SelfC: Self-care/ Pain: Pain and Discomfort/ Cognit: cognition/ InterP: Interpersonal activities/ SleepEng: Sleep and energy/ Affect: affects/ Vision: vision*

Uniq: uniqueness.

Factor loadings lower than 0.2 are omitted.

Larger factor loadings are marked with bold.

Italic numbers show cross-loadings.

Table S5.12. Exploratory Factor Analysis of the Health State Description Questionnaire. South Africa. SAGE study. wave 1.

	SelfC-Gen-Mob	InterP	Pain-Gen	SleepEng	Vision	Affect	Cognit	MobSelfC
	WLS5	**WLS2**	**WLS1**	**WLS7**	**WLS3**	**WLS4**	**WLS6**	**Uniq**
q2000			0.49	*0.20*				0.41
q2001	*0.41*		*0.34*				*0.22*	0.27
q2002	**0.59**		*0.30*					0.18
q2003	0.49		*0.31*					0.28
q2004	**0.88**							0.07
q2005	**0.86**							0.08
q2006	**0.82**							0.19
q2007			**0.91**					0.06
q2008			**0.89**					0.05
q2009			**0.83**					0.08
q2010							**0.75**	0.17
q2011							**0.79**	0.16
q2012		**0.76**						0.17
q2013		**0.89**						0.15
q2014		**0.87**						0.14
q2015		**0.83**						0.28
q2016				**0.96**				0.11
q2017				**0.90**				0.09
q2018						**0.82**		0.10
q2019						**0.84**		0.12
q2023					**0.88**			0.17
q2024					**0.91**			0.18

n	2,073	RMSEA	0.14
rotation	oblimin	lower	0.14
method	wls	upper	0.15
		confidence	0.9

Factors: *Gen: general health/ Mob: mobility/ SelfC: Self-care/ Pain: Pain and Discomfort/ Cognit: cognition/ InterP: Interpersonal activities/ SleepEng: Sleep and energy/ Affect: affects/ Vision: vision*

Uniq: uniqueness.

Factor loadings lower than 0.2 are omitted.

Larger factor loadings are marked with bold.

Italic numbers show cross-loadings.

Figures S5.13-18. Path-Diagram from the selected confirmatory factor analysis from the Health State Description Questionnaire, in six countries. SAGE study. wave 1.

Figure S5.13. China

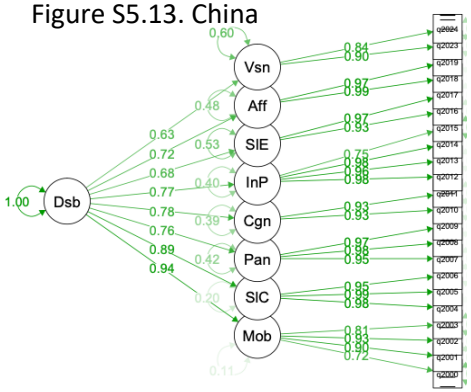


Figure S5.14. Ghana

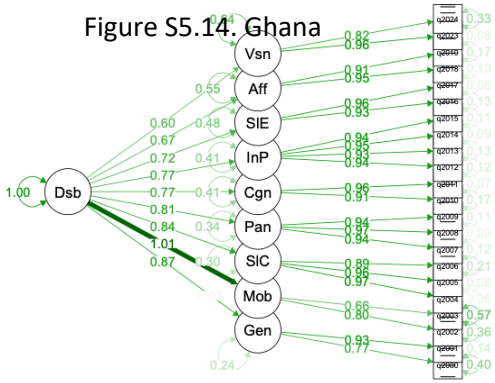


Figure S5.15. India

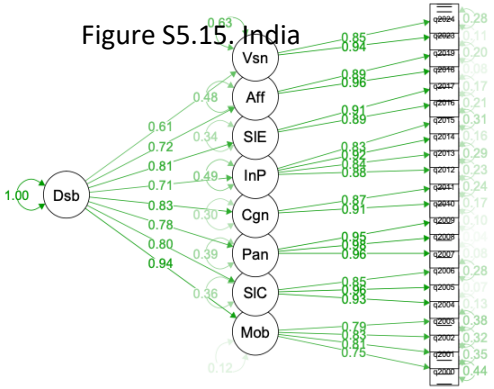


Figure S5.16. Mexico

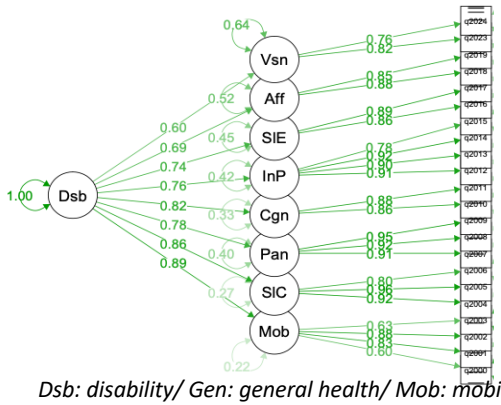


Figure S5.17. Russian Federation

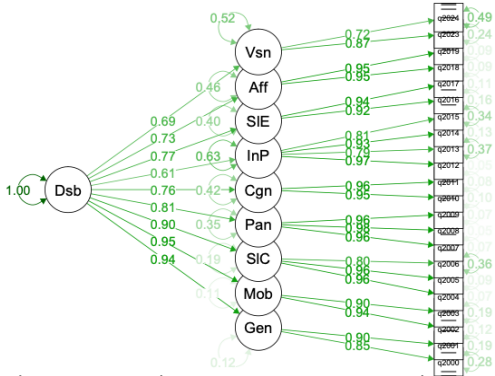
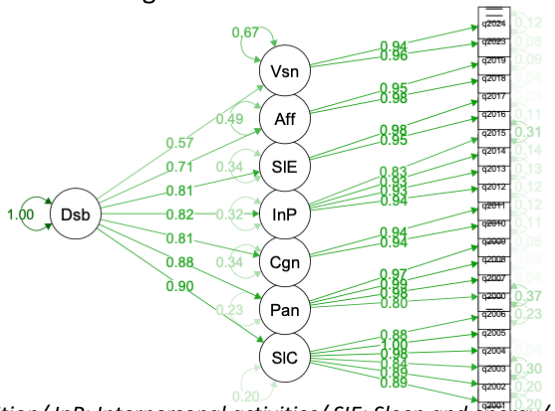


Figure S5.18. South Africa



Dsb: disability/ Gen: general health/ Mob: mobility/ SIC: Self-care/ Pan: Pain and Discomfort/ Cgn: cognition/ InP: Interpersonal activities/ SIE: Sleep and energy/ Aff: affects/ Vsn: vision

Table S5.13. Description of missing data by country in variables using for effective coverage 's measurement. SAGE study. wave 1.

n original	14813	5110	11230	2756	4355	4225	42489
n without missing	13311	4819	10846	2451	3603	3056	38086
% without missing	89.9%	94.3%	96.6%	88.9%	82.7%	72.3%	89.6%

	% of missing data					
	10.1%	5.7%	3.4%	11.1%	17.3%	27.7%

	China	Ghana	India	Mexico	Russia	SouthAfrica
Age	2	2	0	14	0	2
Sex	2	0	0	14	0	2
EdCat	0	33	0	119	3	705
MarEst	16	31	1	119	11	79
Occupation	359	19	1	127	17	136
quintile_c	72	7	71	4	5	24
Zone	0	0	0	0	0	6
Hta	604	68	147	171	94	123
HtaG0	604	68	147	171	94	123
HtaG1	643	68	147	171	94	124
HtaG2	48	0	0	0	0	1
SBP	723	75	157	215	130	131
DBP	723	75	158	215	130	134
Depre	0	0	0	0	0	0
DepreG0	0	0	0	0	0	0
DepreG1	0	0	0	0	0	0
DepreG2	0	0	0	0	0	0
DepreG0b	0	0	0	0	0	0
DepreG1b	0	0	0	0	0	0
DepreG2b	0	0	0	0	0	0
DepreSev	0	0	0	0	0	0
OA	304	18	4	122	43	193
OAG0	304	18	4	122	43	193
OAG1	304	18	4	122	43	193
OAG2	0	0	0	0	0	0
Disab	593	133	156	120	451	349
pweight	19	2	0	23	203	4

	China	Ghana	India	Mexico	Russia	SouthAfrica
Age	0.0%	0.0%	0.0%	0.5%	0.0%	0.0%
Sex	0.0%	0.0%	0.0%	0.5%	0.0%	0.0%
EdCat	0.0%	0.6%	0.0%	4.3%	0.1%	16.7%
MarEst	0.1%	0.6%	0.0%	4.3%	0.3%	1.9%
Occupation	2.4%	0.4%	0.0%	4.6%	0.4%	3.2%
quintile_c	0.5%	0.1%	0.6%	0.1%	0.1%	0.6%
Zone	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
Hta	4.1%	1.3%	1.3%	6.2%	2.2%	2.9%
HtaG0	4.1%	1.3%	1.3%	6.2%	2.2%	2.9%
HtaG1	4.3%	1.3%	1.3%	6.2%	2.2%	2.9%
HtaG2	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%
SBP	4.9%	1.5%	1.4%	7.8%	3.0%	3.1%
DBP	4.9%	1.5%	1.4%	7.8%	3.0%	3.2%
Depre	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
DepreG0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
DepreG1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
DepreG2	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
DepreG0b	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
DepreG1b	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
DepreG2b	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
DepreSev	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
OA	2.1%	0.4%	0.0%	4.4%	1.0%	4.6%
OAG0	2.1%	0.4%	0.0%	4.4%	1.0%	4.6%
OAG1	2.1%	0.4%	0.0%	4.4%	1.0%	4.6%
OAG2	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Disab	4.0%	2.6%	1.4%	4.4%	10.4%	8.3%
pweight	0.1%	0.0%	0.0%	0.8%	4.7%	0.1%

EdCat: education/ MarEst: marital status/ HTA: hypertension/ SBP: systolic blood pressure/ DBP: diastolic blood pressure/ Depre: depression/ DepreSev: depression severity/ OA: osteoarthritis/ Disab: disability/ pweight: sample weight. / Cells in gray shows missing rates higher than 5%.

Table S5.14. Coefficients for disability from regression models by country used to calculate effective coverage of the treatment for depression. SAGE study. wave 1.

	n	13311
	China	
	Coef	**CI**
(Intercept)	5.44	[2.4 - 2.42]
Age2	3.67	[3.3 - 3.27]
Sex	1.53	[0.6 - 0.61]
Zone	1.98	[0.9 - 0.89]
ME1	2.08	[-0.7 - -0.74]
ME2	-0.85	[-4.0 - -4.03]
ME3	-0.02	[-1.2 - -1.19]
Ed1	-1.01	[-2.4 - -2.44]
Ed2	-2.54	[-3.8 - -3.84]
Ed3	-3.07	[-5.4 - -5.40]
Q1	-2.65	[-4.2 - -4.21]
Q2	-3.55	[-5.1 - -5.13]
Q3	-4.67	[-6.1 - -6.13]
Q4	-5.85	[-7.4 - -7.41]
Oc1	2.17	[0.3 - 0.28]
Oc2	4.59	[3.4 - 3.39]
Hta	0.61	[-0.3 - -0.33]
OA	6.30	[4.6 - 4.65]
DepreG1b	17.21	[0.4 - 0.42]
DepreG2b	25.84	[-30.2 - -30.24]
DepreSev	2.43	[1.8 - 1.80]
Age2:DepreG1b	-2.60	[-5.7 - -5.66]
Age2:DepreG2b	-2.38	[-18.9 - -18.87]
Sex:DepreG1b	-2.94	[-14.1 - -14.14]
Sex:DepreG2b	6.20	[-22.8 - -22.79]
DepreG2b:DepreSev	-2.26	[-10.4 - -10.41]

	n	4819
	Ghana	
	Coef	**CI**
(Intercept)	-0.88	[-4.7 - -4.70]
Age2	4.20	[3.7 - 3.68]
Sex	1.63	[-0.1 - -0.06]
Zone	3.10	[1.5 - 1.46]
ME1	4.49	[1.3 - 1.33]
ME2	1.31	[-1.7 - -1.70]
ME3	3.59	[1.5 - 1.46]
Ed1	-0.70	[-2.9 - -2.89]
Ed2	-0.64	[-2.6 - -2.59]
Ed3	-0.49	[-4.0 - -3.96]
Q1	0.64	[-1.9 - -1.85]
Q2	0.50	[-2.2 - -2.22]
Q3	-3.44	[-6.0 - -6.04]
Q4	-2.61	[-5.3 - -5.29]
Oc1	-2.87	[-6.9 - -6.86]
Oc2	5.85	[3.7 - 3.71]
Hta	-0.20	[-1.7 - -1.74]
OA	11.09	[8.9 - 8.86]
DepreG1b	-38.24	[-63.1 - -63.11]
DepreG2b	-9.07	[-67.3 - -67.32]
DepreSev	6.65	[3.7 - 3.74]
Age2:DepreG2b	-0.96	[-4.9 - -4.93]
Sex:DepreG2b	5.16	[-3.5 - -3.52]
DepreG2b:DepreSev	1.35	[-5.7 - -5.65]
Age2:DepreG1b	3.74	[-0.1 - -0.14]
Age2:DepreSev	-0.78	[-1.2 - -1.24]
Sex:DepreSev	-0.56	[-1.3 - -1.30]
Q1:DepreG1b	2.25	[-4.2 - -4.22]
Q2:DepreG1b	5.24	[-0.9 - -0.86]
Q3:DepreG1b	7.34	[1.3 - 1.33]
Q4:DepreG1b	8.84	[1.5 - 1.50]
Q1:DepreG2b	-9.20	[-17.2 - -17.22]
Q3:DepreG2b	5.65	[-6.0 - -6.00]
Ed2:DepreG1b	1.57	[-4.0 - -4.00]
ME1:DepreG1b	-20.25	[-27.5 - -27.53]
ME2:DepreG1b	2.89	[-2.8 - -2.78]
ME3:DepreG1b	3.47	[-3.9 - -3.89]
Oc2:DepreG2b	5.55	[-0.8 - -0.84]
Oc2:DepreG1b	3.40	[-0.9 - -0.94]

	n	10846
	India	
	Coef	**CI**
(Intercept)	7.44	[5.4 - 5.36]
Age2	4.00	[3.7 - 3.70]
Sex	4.00	[3.0 - 3.01]
Zone	1.63	[0.7 - 0.67]
ME1	-0.23	[-1.7 - -1.66]
ME2	-0.32	[-3.9 - -3.87]
ME3	1.71	[0.4 - 0.43]
Ed1	-0.39	[-1.5 - -1.51]
Ed2	-2.71	[-3.7 - -3.70]
Ed3	-4.89	[-6.5 - -6.54]
Q1	-0.66	[-1.9 - -1.89]
Q2	-1.58	[-2.8 - -2.82]
Q3	-2.34	[-3.6 - -3.62]
Q4	-4.17	[-5.4 - -5.43]
Oc1	0.66	[-0.3 - -0.31]
Oc2	1.64	[0.7 - 0.68]
Hta	2.00	[1.0 - 0.98]
OA	5.42	[4.2 - 4.19]
DepreG1b	-6.70	[-15.9 - -15.88]
DepreG2b	26.89	[-12.4 - -12.44]
DepreSev	3.22	[2.4 - 2.36]
Age2:DepreG1b	0.63	[-0.7 - -0.70]
Age2:DepreG2b	0.39	[-1.9 - -1.90]
Sex:DepreG2b	5.62	[-6.3 - -6.28]
DepreG2b:DepreSev	-3.24	[-7.2 - -7.16]
Age2:DepreSev	-0.23	[-0.4 - -0.37]
Sex:DepreSev	-0.26	[-0.6 - -0.63]
Q1:DepreG1b	1.69	[-1.9 - -1.86]
Q2:DepreG1b	-1.44	[-4.6 - -4.56]
Q3:DepreG1b	-2.37	[-5.8 - -5.75]
Q1:DepreG2b	-5.06	[-13.4 - -13.35]
Q3:DepreG2b	-8.88	[-17.4 - -17.44]
Q4:DepreG2b	-12.20	[-24.7 - -24.73]
Ed2:DepreG1b	4.22	[0.5 - 0.47]
Ed3:DepreG1b	8.13	[3.4 - 3.38]
Ed2:DepreG2b	11.44	[-1.0 - -0.99]
ME3:DepreG1b	-1.30	[-4.2 - -4.16]
ME1:DepreG2b	-8.43	[-24.5 - -24.51]
ME2:DepreG2b	-31.76	[-48.7 - -48.71]
Oc1:DepreG2b	-0.12	[-12.9 - -12.85]
Oc2:DepreG2b	9.08	[-1.0 - -1.03]
Oc2:DepreG1b	4.79	[1.7 - 1.68]
Zone:DepreG1b	0.86	[-2.9 - -2.88]

- *Age2: age*
- *ME: marital status (1: never married; 2:separated or divorced; 3: widowed; reference: married)*
- *Ed: education (1: primary completed; 2 secondary completed; 3 college or university completed; reference: less than primary)*
- *Q: quintil of income*
- *Oc: occupation (1: never worked/ homemaker; 2: not working; reference: working)*
- *Hta: hypertension*
- *OA: osteoarthrosis*
- *DepreG1b: depression without treatment*
- *DepreG2b: depression with treatment*
- *DepreSev: Severity of the depressive disorder*

Zone:DepreG2b	-7.07	[-12.6 - -12.64]

Table S5.14. Continuation

	n	2451
	Mexico	
	Coef	**CI**
(Intercept)	12.46	[6.5 - 6.55]
Age2	2.29	[1.5 - 1.51]
Sex	3.43	[0.8 - 0.79]
Zone	2.40	[-0.4 - -0.44]
ME1	0.12	[-2.8 - -2.75]
ME2	4.43	[-0.5 - -0.49]
ME3	-1.19	[-5.1 - -5.06]
Ed1	-1.71	[-4.9 - -4.88]
Ed2	-4.32	[-7.4 - -7.36]
Ed3	-3.02	[-7.1 - -7.07]
Q1	-2.56	[-6.7 - -6.65]
Q2	-2.26	[-6.1 - -6.10]
Q3	-6.17	[-9.7 - -9.66]
Q4	-5.78	[-9.6 - -9.61]
Oc1	0.12	[-2.7 - -2.75]
Oc2	3.05	[0.3 - 0.29]
Hta	-0.43	[-2.8 - -2.78]
OA	3.73	[-0.3 - -0.32]
DepreG1b	-23.93	[-52.1 - -52.07]
DepreG2b	-1.88	[-28.7 - -28.69]
DepreSev	4.68	[1.9 - 1.88]
Age2:DepreG1b	2.91	[-1.4 - -1.39]
Age2:DepreG2b	-0.14	[-3.8 - -3.77]
Sex:DepreG1b	16.71	[-0.1 - -0.08]
Sex:DepreG2b	9.52	[-5.1 - -5.06]
DepreG2b:DepreSev	-0.40	[-2.5 - -2.53]
Age2:DepreSev	-0.17	[-0.7 - -0.67]
Sex:DepreSev	-1.75	[-3.6 - -3.60]
Q1:DepreG1b	-9.90	[-17.6 - -17.60]
Q2:DepreG1b	-11.79	[-17.5 - -17.49]
Q4:DepreG1b	-7.59	[-15.0 - -15.03]
Q3:DepreG2b	9.25	[4.8 - 4.78]
Q4:DepreG2b	11.92	[6.3 - 6.25]
Ed1:DepreG1b	-3.77	[-9.3 - -9.29]
Ed2:DepreG1b	6.79	[-0.8 - -0.84]
Ed1:DepreG2b	-6.02	[-13.0 - -13.00]
Ed2:DepreG2b	-7.07	[-11.2 - -11.21]
ME2:DepreG1b	-15.78	[-24.3 - -24.33]
ME3:DepreG1b	-6.00	[-13.2 - -13.21]

	n	3603
	Russia	
	Coef	**CI**
(Intercept)	8.40	[0.3 - 0.34]
Age2	4.11	[3.3 - 3.25]
Sex	0.00	[-3.5 - -3.55]
Zone	4.43	[1.8 - 1.83]
ME1	0.71	[-3.7 - -3.66]
ME2	0.97	[-2.2 - -2.21]
ME3	0.55	[-1.9 - -1.90]
Ed1	-6.33	[-11.8 - -11.78]
Ed2	-8.91	[-14.2 - -14.20]
Ed3	-7.92	[-13.8 - -13.84]
Q1	0.36	[-2.5 - -2.48]
Q2	0.22	[-2.8 - -2.76]
Q3	-3.22	[-7.0 - -7.04]
Q4	-2.24	[-5.8 - -5.83]
Oc1	4.89	[-3.1 - -3.10]
Oc2	6.47	[4.3 - 4.32]
Hta	1.61	[-1.9 - -1.88]
OA	4.63	[2.4 - 2.37]
DepreG1b	-45.51	[-71.0 - -71.04]
DepreG2b	-34.53	[-66.9 - -66.89]
DepreSev	5.48	[2.4 - 2.45]
Age2:DepreG1b	4.64	[0.4 - 0.44]
Age2:DepreG2b	5.72	[0.8 - 0.80]
Sex:DepreG1b	15.94	[2.4 - 2.44]
Sex:DepreG2b	16.03	[1.2 - 1.21]
DepreG2b:DepreSev	-2.18	[-4.8 - -4.82]
Age2:DepreSev	-0.24	[-0.8 - -0.78]
Sex:DepreSev	-1.73	[-3.5 - -3.46]
Q3:DepreG1b	2.15	[-4.7 - -4.68]
Q1:DepreG2b	-9.84	[-19.9 - -19.90]
Ed1:DepreG1b	11.46	[0.7 - 0.67]
Ed2:DepreG1b	6.35	[0.6 - 0.59]
ME1:DepreG1b	-5.32	[-14.6 - -14.60]
ME3:DepreG1b	-7.06	[-14.1 - -14.11]
ME2:DepreG2b	-5.94	[-14.2 - -14.20]
ME3:DepreG2b	-12.82	[-24.4 - -24.40]
Oc2:DepreG2b	5.48	[-2.8 - -2.77]
Oc1:DepreG1b	-6.42	[-19.7 - -19.70]
Oc2:DepreG1b	-2.32	[-9.6 - -9.61]

	n	3056
	South Africa	
	Coef	**CI**
(Intercept)	0.97	[-9.1 - -9.12]
Age2	3.83	[2.6 - 2.56]
Sex	-0.79	[-6.8 - -6.77]
Zone	-5.99	[-11.1 - -11.06]
ME1	3.25	[-3.4 - -3.40]
ME2	-0.31	[-6.6 - -6.61]
ME3	1.58	[-2.9 - -2.89]
Ed1	-1.69	[-6.3 - -6.25]
Ed2	-0.89	[-6.3 - -6.30]
Ed3	1.37	[-4.9 - -4.91]
Q1	0.37	[-5.8 - -5.81]
Q2	-1.21	[-9.6 - -9.61]
Q3	-3.52	[-8.2 - -8.18]
Q4	-5.56	[-11.6 - -11.59]
Oc1	9.26	[3.3 - 3.33]
Oc2	6.77	[1.8 - 1.78]
Hta	4.33	[-0.5 - -0.51]
OA	8.73	[4.3 - 4.33]
DepreG1b	-13.01	[-65.3 - -65.33]
DepreG2b	-47.90	[-150.9 - -150.87]
DepreSev	6.87	[0.2 - 0.18]
Age2:DepreG1b	2.68	[-6.1 - -6.05]
Age2:DepreG2b	7.08	[-6.2 - -6.24]
Sex:DepreG1b	13.88	[-7.5 - -7.51]
Sex:DepreG2b	24.26	[1.3 - 1.34]
DepreG2b:DepreSev	2.28	[-4.5 - -4.48]
Age2:DepreSev	-0.54	[-1.7 - -1.67]
Sex:DepreSev	-1.24	[-3.6 - -3.62]
Q1:DepreG1b	-15.96	[-26.4 - -26.41]
Q2:DepreG1b	-7.08	[-17.9 - -17.92]
Q4:DepreG1b	-12.42	[-22.1 - -22.15]
Q1:DepreG2b	-14.52	[-27.7 - -27.74]
Q2:DepreG2b	-18.34	[-34.5 - -34.55]
Q4:DepreG2b	-7.46	[-21.7 - -21.73]
Ed3:DepreG1b	-15.75	[-36.6 - -36.59]
Ed1:DepreG2b	-11.14	[-27.0 - -26.96]
Ed2:DepreG2b	-12.86	[-31.5 - -31.53]
ME2:DepreG1b	10.42	[0.3 - 0.28]
ME3:DepreG1b	7.29	[-2.4 - -2.37]

Oc1:DepreG2b	-8.76	[-15.0 - -14.99]
Oc2:DepreG2b	-9.23	[-14.9 - -14.90]
Oc1:DepreG1b	-4.46	[-11.1 - -11.07]
Oc2:DepreG1b	-5.36	[-13.8 - -13.77]
Zone:DepreG1b	-9.58	[-18.0 - -17.99]

ME3:DepreG2b	-10.21	[-25.3 - -25.32]
Oc1:DepreG2b	-27.14	[-42.2 - -42.20]
Oc2:DepreG2b	-21.60	[-35.2 - -35.23]
Oc1:DepreG1b	-26.99	[-51.7 - -51.74]
Oc2:DepreG1b	-17.18	[-38.7 - -38.72]

Figure S5.19. Histograms of observed health gain attributable to healthcare interventions addressed to depressive disorder by country. SAGE study. wave 1.

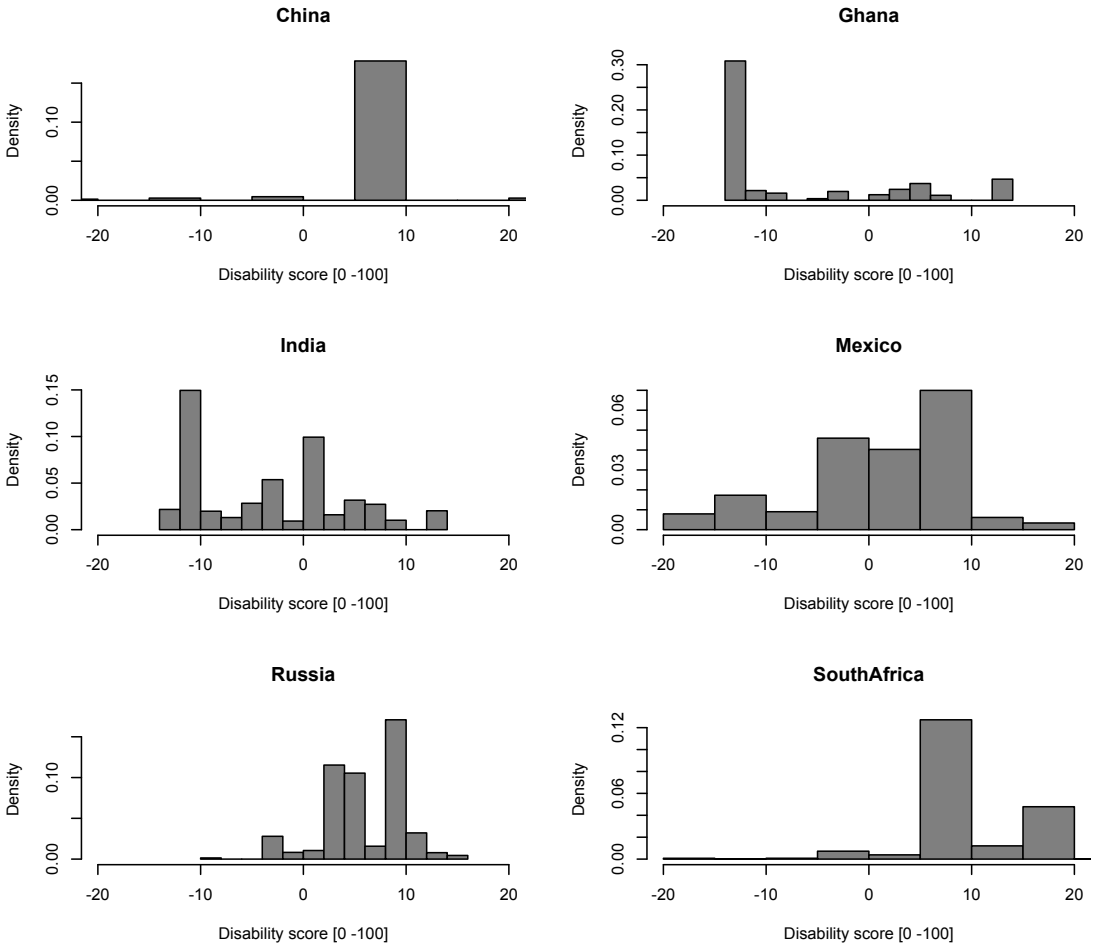


Table S5.15. Results of effective coverage of the healthcare for depression, by country, assuming effective coverage as an individual attribute. SAGE study, wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	23.9	[14.9 – 33.0]	71.3	[-359.6 - 502.2]	-27.6	[-50.6 - -4.6]
Quality (%)	82.4	[45.3 - 119.5]	-126.1	[-243.9 - -8.4]	-54.9	[-103.7 - -6.2]
relative - effective coverage (%)	3.4	[-2.1 – 9.0]	5.4	[-24.9 - 35.7]	-1.5	[-3.1 - 0.1]
absolute - effective coverage (%)	0.0	[0 .0- 0.1]	0.2	[-1.1 - 1.5]	-0.1	[-0.2 – 0.0]
Effective coverage (%)	11.8	[-10.9 - 34.5]	-9.5	[-28.0 – 9.0]	-3.0	[-6.6 - 0.6]
RR HG-average	1.31	[1.17 - 1.49]	3.49	[0.22 - Inf]	0.78	[0.66 - 0.96]
RR HG-max	1.41	[1.22 - 1.68]	Inf	[0.32 - Inf]	2.13	[1.71 - 2.84]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Afriaca	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	-4.2	[-117.1 - 108.6]	27.8	[20.1 - 35.6]	72.6	[-24.7 - 169.9]
Quality (%)	8.2	[-48.4 - 64.8]	63.6	[40.6 - 86.7]	53.9	[34.1 - 73.7]
relative - effective coverage (%)	-1.0	[-28.4 - 26.3]	6.8	[1.2 - 12.4]	55.1	[-11.8 - 122]
absolute - effective coverage (%)	-0.1	[-1.7 - 1.6]	0.2	[0.0 - 0.4]	2.7	[-0.9 - 6.3]
Effective coverage (%)	2.0	[-12.1 - 16.1]	15.6	[2.1 - 29.2]	40.9	[24.4 - 57.4]
RR HG-average	0.96	[0.46 - Inf]	1.39	[1.25 - 1.55]	3.65	[0.8 - Inf]
RR HG-max	3.70	[1.19 - Inf]	2.02	[1.69 - 2.5]	Inf	[1.02 - Inf]

Table S5.16. Coefficients for disability from regression models by country used to calculate effective coverage of the treatment for hypertension. SAGE study. wave 1.

	n	13311
	China	
	Coef	**CI**
(Intercept)	8.58	[4.9 - 4.88]
Age2	3.18	[2.7 - 2.70]
Sex	1.06	[-0.1 - -0.05]
Zone	1.63	[0.2 - 0.24]
ME1	1.82	[-0.9 - -0.90]
ME2	-0.42	[-4.4 - -4.37]
ME3	0.03	[-1.1 - -1.13]
Ed1	-2.09	[-4.1 - -4.07]
Ed2	-3.97	[-5.8 - -5.78]
Ed3	-4.86	[-7.8 - -7.76]
Q1	-1.55	[-3.4 - -3.39]
Q2	-1.84	[-3.8 - -3.84]
Q3	-3.32	[-5.0 - -5.01]
Q4	-4.49	[-6.3 - -6.33]
Oc1	1.29	[-0.9 - -0.88]
Oc2	4.44	[3.2 - 3.22]
Depre	20.33	[15.5 - 15.51]
OA	6.06	[4.2 - 4.19]
HtaG1	-3.47	[-9.9 - -9.90]
HtaG2	-2.88	[-10.5 - -10.46]
Age2:HtaG1	0.87	[0.1 - 0.13]
Age2:HtaG2	0.69	[-0.3 - -0.34]
Sex:HtaG1	1.45	[-0.4 - -0.40]
Q1:HtaG1	-4.57	[-8.0 - -7.97]
Q2:HtaG1	-6.32	[-9.8 - -9.77]
Q3:HtaG1	-6.27	[-9.3 - -9.32]
Q4:HtaG1	-6.35	[-9.8 - -9.79]
Q2:HtaG2	-2.37	[-5.3 - -5.25]
Q4:HtaG2	-3.54	[-6.9 - -6.87]
Ed1:HtaG1	2.52	[-0.4 - -0.41]
Ed2:HtaG1	3.86	[1.2 - 1.19]
Ed3:HtaG1	2.47	[-2.3 - -2.29]
Ed1:HtaG2	1.59	[-2.1 - -2.15]
Ed2:HtaG2	2.97	[-0.1 - -0.06]
Ed3:HtaG2	9.07	[1.8 - 1.81]
ME2:HtaG1	-3.05	[-8.7 - -8.74]
ME2:HtaG2	5.72	[0.4 - 0.43]
Oc1:HtaG1	3.92	[-0.8 - -0.81]
Zone:HtaG1	1.74	[-0.4 - -0.45]
Zone:HtaG2	3.02	[0.2 - 0.24]
OA:HtaG1	3.55	[0.0 - 0.05]

	n	4819
	Ghana	
	Coef	**CI**
(Intercept)	-1.11	[-5.3 - -5.27]
Age2	4.22	[3.6 - 3.64]
Sex	1.91	[0.2 - 0.22]
Zone	3.88	[1.9 - 1.85]
ME1	2.46	[-1.3 - -1.29]
ME2	3.54	[-0.1 - -0.14]
ME3	4.14	[1.8 - 1.81]
Ed1	-0.09	[-2.3 - -2.27]
Ed2	0.19	[-2.1 - -2.09]
Ed3	0.71	[-2.6 - -2.58]
Q1	0.46	[-2.3 - -2.26]
Q2	-0.68	[-3.8 - -3.82]
Q3	-3.87	[-6.7 - -6.73]
Q4	-1.66	[-4.3 - -4.26]
Oc1	-3.69	[-7.6 - -7.64]
Oc2	4.25	[1.7 - 1.69]
Depre	5.95	[2.7 - 2.66]
OA	9.66	[7.1 - 7.07]
HtaG1	0.62	[-5.2 - -5.21]
HtaG2	20.22	[9.2 - 9.16]
Age2:HtaG1	-0.79	[-1.7 - -1.70]
Age2:HtaG2	-2.33	[-3.9 - -3.95]
Q1:HtaG1	3.70	[-0.8 - -0.82]
Q2:HtaG1	5.40	[0.8 - 0.76]
Q3:HtaG1	3.40	[-0.4 - -0.41]
Q2:HtaG2	6.06	[1.0 - 1.03]
Ed2:HtaG1	-1.59	[-5.0 - -5.02]
Ed2:HtaG2	-7.94	[-12.8 - -12.81]
ME1:HtaG1	7.24	[0.7 - 0.73]
ME2:HtaG1	-3.21	[-9.0 - -8.99]
ME1:HtaG2	-10.68	[-17.0 - -16.96]
ME2:HtaG2	-12.18	[-18.6 - -18.63]
ME3:HtaG2	-6.03	[-9.9 - -9.86]
Oc1:HtaG2	17.90	[11.3 - 11.27]
Oc2:HtaG1	6.31	[2.6 - 2.57]
Oc2:HtaG2	4.72	[-0.1 - -0.11]
Zone:HtaG1	-1.19	[-4.5 - -4.48]
Zone:HtaG2	-3.17	[-7.6 - -7.61]
Depre:HtaG1	-4.24	[-8.6 - -8.62]
OA:HtaG1	3.06	[-1.4 - -1.44]
OA:HtaG2	3.07	[-1.3 - -1.29]

	n	10846
	India	
	Coef	**CI**
(Intercept)	7.80	[5.6 - 5.62]
Age2	4.06	[3.8 - 3.77]
Sex	4.19	[3.2 - 3.15]
Zone	1.95	[0.9 - 0.93]
ME1	-0.82	[-2.3 - -2.34]
ME2	-2.05	[-6.4 - -6.43]
ME3	1.47	[0.2 - 0.21]
Ed1	0.05	[-1.1 - -1.14]
Ed2	-1.97	[-3.1 - -3.07]
Ed3	-4.79	[-6.5 - -6.45]
Q1	-0.57	[-1.8 - -1.81]
Q2	-1.60	[-2.9 - -2.86]
Q3	-2.84	[-4.1 - -4.14]
Q4	-4.58	[-5.9 - -5.95]
Oc1	0.64	[-0.4 - -0.37]
Oc2	1.69	[0.6 - 0.59]
Depre	12.52	[10.8 - 10.83]
OA	6.47	[5.2 - 5.19]
HtaG1	6.83	[1.5 - 1.53]
HtaG2	4.06	[-0.1 - -0.07]
Age2:HtaG1	-0.40	[-1.2 - -1.19]
Sex:HtaG1	-4.05	[-7.2 - -7.17]
Sex:HtaG2	1.40	[-1.8 - -1.84]
Q2:HtaG1	-2.06	[-5.2 - -5.22]
Q4:HtaG1	-1.40	[-4.0 - -4.03]
Q1:HtaG2	-2.21	[-7.1 - -7.10]
Q3:HtaG2	-2.11	[-7.2 - -7.20]
Q4:HtaG2	-3.73	[-8.1 - -8.12]
Ed1:HtaG1	-2.83	[-6.7 - -6.73]
Ed2:HtaG1	-2.29	[-5.1 - -5.07]
Ed1:HtaG2	-1.60	[-5.4 - -5.44]
Ed2:HtaG2	-0.88	[-4.8 - -4.80]
ME1:HtaG1	4.92	[-2.1 - -2.06]
ME2:HtaG1	9.79	[2.4 - 2.37]
ME3:HtaG1	-0.87	[-4.1 - -4.07]
ME2:HtaG2	3.15	[-5.7 - -5.68]
Oc2:HtaG2	3.11	[-0.4 - -0.37]
Oc1:HtaG1	2.33	[-1.2 - -1.23]
Oc2:HtaG1	2.76	[-0.2 - -0.16]
Zone:HtaG1	-1.47	[-4.4 - -4.36]
Depre:HtaG1	-1.25	[-4.5 - -4.52]

Table S5.16. Continuation

	n	2451
	Mexico	
	Coef	**CI**
(Intercept)	13.43	[6.4 - 6.41]
Age2	2.82	[1.8 - 1.79]
Sex	4.45	[1.9 - 1.88]
Zone	0.24	[-3.0 - -3.01]
ME1	0.98	[-1.8 - -1.81]
ME2	2.73	[-1.0 - -1.04]
ME3	-0.95	[-3.8 - -3.76]
Ed1	-2.39	[-5.6 - -5.55]
Ed2	-4.76	[-7.9 - -7.89]
Ed3	-2.38	[-6.6 - -6.56]
Q1	-1.53	[-6.1 - -6.06]
Q2	-4.18	[-8.1 - -8.13]
Q3	-7.24	[-11.0 - -10.96]
Q4	-8.16	[-12.0 - -11.99]
Oc1	-0.29	[-3.0 - -2.99]
Oc2	0.24	[-3.1 - -3.07]
Depre	9.34	[5.7 - 5.70]
OA	7.94	[3.6 - 3.62]
HtaG1	3.71	[-3.6 - -3.57]
HtaG2	4.73	[-7.2 - -7.22]
Age2:HtaG1	-0.75	[-2.1 - -2.08]
Age2:HtaG2	-1.25	[-3.0 - -3.01]
Sex:HtaG1	-3.03	[-6.9 - -6.90]
Q1:HtaG1	-10.34	[-15.8 - -15.82]
Q3:HtaG1	-3.23	[-7.7 - -7.67]
Q2:HtaG2	4.05	[-0.9 - -0.90]
Q3:HtaG2	6.75	[1.4 - 1.43]
Q4:HtaG2	8.62	[2.9 - 2.87]
Ed3:HtaG1	-7.08	[-13.8 - -13.79]
Ed3:HtaG2	-8.21	[-15.2 - -15.22]
ME1:HtaG2	-6.50	[-11.7 - -11.73]
Oc2:HtaG2	2.45	[-2.0 - -2.01]
Oc2:HtaG1	5.34	[0.5 - 0.50]
Zone:HtaG1	4.05	[-0.8 - -0.75]
OA:HtaG2	-3.77	[-9.3 - -9.35]

	n	3603
	Russia	
	Coef	**CI**
(Intercept)	7.78	[-0.3 - -0.34]
Age2	4.07	[3.1 - 3.10]
Sex	-0.02	[-3.4 - -3.41]
Zone	5.52	[2.5 - 2.53]
ME1	0.21	[-4.7 - -4.66]
ME2	1.15	[-3.1 - -3.06]
ME3	-2.24	[-7.0 - -6.95]
Ed1	-4.23	[-8.7 - -8.69]
Ed2	-8.31	[-12.4 - -12.41]
Ed3	-8.39	[-13.5 - -13.51]
Q1	5.43	[1.5 - 1.54]
Q2	1.14	[-1.8 - -1.78]
Q3	-0.02	[-3.4 - -3.45]
Q4	-0.90	[-4.4 - -4.45]
Oc1	5.52	[-1.3 - -1.29]
Oc2	6.37	[3.9 - 3.87]
Depre	13.01	[8.4 - 8.41]
OA	8.52	[4.5 - 4.49]
HtaG1	10.70	[-9.9 - -9.87]
HtaG2	20.67	[10.6 - 10.60]
Age2:HtaG1	0.92	[-1.0 - -1.01]
Age2:HtaG2	-2.18	[-3.8 - -3.80]
Sex:HtaG1	-4.26	[-9.7 - -9.68]
Q1:HtaG1	-8.56	[-14.2 - -14.23]
Q2:HtaG1	-3.98	[-8.4 - -8.44]
Q3:HtaG1	-9.64	[-16.3 - -16.27]
Q1:HtaG2	-7.41	[-11.4 - -11.38]
Ed1:HtaG1	-13.05	[-30.1 - -30.11]
Ed2:HtaG1	-10.80	[-27.2 - -27.24]
Ed3:HtaG1	-9.60	[-26.5 - -26.53]
ME1:HtaG1	8.37	[-1.7 - -1.69]
ME2:HtaG1	9.10	[1.1 - 1.10]
ME3:HtaG1	10.15	[3.6 - 3.59]
ME2:HtaG2	-2.59	[-8.3 - -8.33]
ME3:HtaG2	3.20	[-1.8 - -1.76]
Oc2:HtaG2	1.77	[-1.9 - -1.91]
Oc2:HtaG1	-3.58	[-8.6 - -8.64]
Zone:HtaG2	-7.54	[-11.3 - -11.32]
OA:HtaG1	-3.08	[-8.5 - -8.45]
Depre:HtaG2	-4.08	[-9.2 - -9.18]
OA:HtaG2	-4.77	[-9.3 - -9.32]

	n	2596
	South Africa	
	Coef	**CI**
(Intercept)	-5.83	[-20.3 - -20.29]
Age2	5.49	[3.3 - 3.26]
Sex	2.30	[-3.3 - -3.31]
Zone	-5.02	[-9.8 - -9.81]
ME1	3.86	[-2.9 - -2.94]
ME2	1.98	[-4.1 - -4.15]
ME3	1.13	[-3.5 - -3.46]
Ed1	-1.64	[-6.3 - -6.27]
Ed2	-1.31	[-6.4 - -6.36]
Ed3	0.82	[-5.0 - -5.00]
Q1	-0.34	[-6.3 - -6.33]
Q2	-2.49	[-10.3 - -10.27]
Q3	-4.02	[-8.6 - -8.63]
Q4	-5.77	[-11.1 - -11.15]
Oc1	6.47	[0.4 - 0.44]
Oc2	6.67	[1.7 - 1.68]
Depre	15.23	[8.6 - 8.64]
OA	9.12	[4.8 - 4.85]
HtaG1	17.25	[-1.6 - -1.56]
HtaG2	25.60	[11.4 - 11.42]
Age2:HtaG1	-2.68	[-5.8 - -5.82]
Age2:HtaG2	-3.85	[-6.3 - -6.34]
Sex:HtaG1	-6.47	[-14.2 - -14.25]

- *Age2: age*
- *ME: marital status (1: never married; 2:separated or divorced; 3: widowed; reference: married)*
- *Ed: education (1: primary completed; 2 secondary completed; 3 college or university completed; reference: less than primary)*
- *Q: quintil of income*
- *Oc: occupation (1: never worked/ homemaker; 2: not working; reference: working)*
- *Depre: depression*
- *OA: osteoarthrosis*
- *HtaG1: hypertension without treatment*
- *HtaG2: hypertension with treatment*

Figure S5.20. Histograms of observed health gain attributable to healthcare interventions addressed to hypertension by country. SAGE study. wave 1.

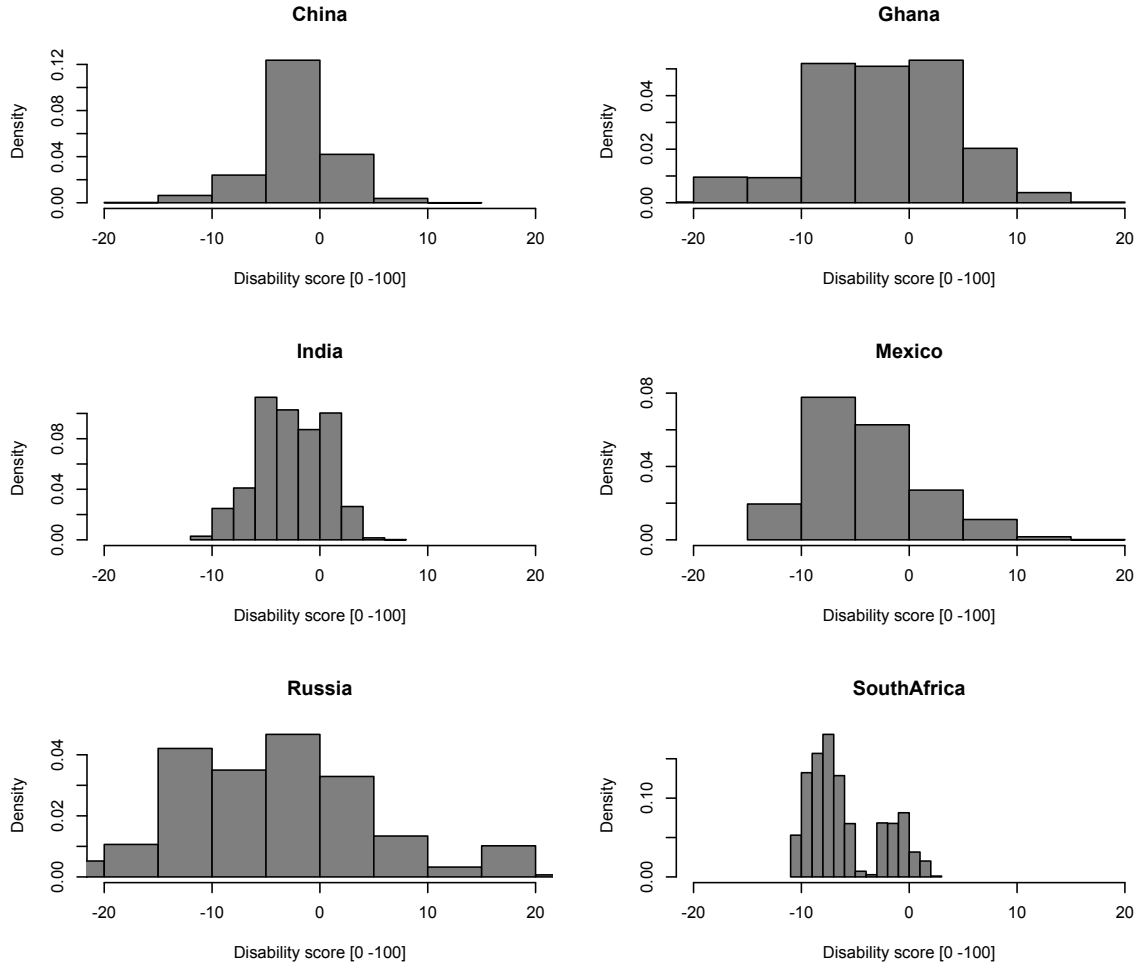


Table S5.17. Results of effective coverage of the healthcare for hypertension, by country, assuming effective coverage as an individual attribute. SAGE study, wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	33.3	[-266.2 - 332.8]	189.1	[-189.5 - 567.8]	5.7	[-235.4 - 246.8]
Quality (%)	-105.8	[-122.1 - -89.5]	-41.3	[-69.5 - -13.1]	-160.2	[-186.5 - -133.8]
relative - effective coverage (%)	8.1	[-65.0 - 81.2]	24.0	[-24.1 - 72.1]	1.3	[-55.2 - 57.9]
absolute - effective coverage (%)	2.7	[-21.9 - 27.4]	8.4	[-8.4 - 25.2]	0.2	[-8.6 – 9.0]
Effective coverage (%)	-25.8	[-31.0 - -20.6]	-5.2	[-9.3 - -1.2]	-37.6	[-44.6 - -30.6]
RR HG-average	1.50	[0.27 - Inf]	Inf	[0.35 - Inf]	1.06	[0.30 - Inf]
RR HG-max	1.43	[0.50 - Inf]	0.33	[0.17 - 3.23]	1.09	[0.52 - Inf]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	5060.2	[-4682.9 - 14803.2]	950.8	[-644.4 - 2545.9]	-544.9	[-1609.9 - 520.2]
Quality (%)	-143.7	[-168.6 - -118.7]	-62.8	[-84.7 - -40.8]	797.4	[721 - 873.9]
relative - effective coverage (%)	1576.2	[-1517.8 - 4670.3]	538.2	[-374.6 - 1451.1]	-95.2	[-287 - 96.5]
absolute - effective coverage (%)	393.3	[-383.5 - 1170.2]	185.2	[-131.8 - 502.1]	-44.6	[-134.5 - 45.3]
Effective coverage (%)	-44.8	[-57.5 - -32.0]	-35.5	[-48.8 - -22.3]	139.4	[99 - 179.8]
RR HG-average	Inf	[0.02 - Inf]	Inf	[0.13 - Inf]	0.16	[0.06 - Inf]
RR HG-max	0.02	[0.01 - Inf]	0.14	[0.05 - Inf]	0.67	[0.44 - 1.48]

Table S5.18. Coefficients for disability from regression models by country used to calculate effective coverage of the treatment for osteoarthritis. SAGE study. wave 1.

	n	13311
	China	
	Coef	**CI**
(Intercept)	6.68	[3.6 - 3.60]
Age2	3.64	[3.2 - 3.23]
Sex	1.46	[0.5 - 0.53]
Zone	2.22	[1.1 - 1.15]
ME1	2.08	[-0.7 - -0.75]
ME2	-1.01	[-4.2 - -4.16]
ME3	0.67	[-0.5 - -0.48]
Ed1	-1.51	[-3.0 - -2.96]
Ed2	-2.54	[-3.9 - -3.93]
Ed3	-3.22	[-5.6 - -5.64]
Q1	-2.86	[-4.5 - -4.53]
Q2	-3.80	[-5.4 - -5.38]
Q3	-5.25	[-6.7 - -6.74]
Q4	-6.79	[-8.4 - -8.39]
Oc1	2.28	[0.3 - 0.30]
Oc2	4.65	[3.4 - 3.43]
Depre	22.54	[18.1 - 18.14]
Hta	0.29	[-0.7 - -0.70]
OAG1	-4.35	[-15.6 - -15.58]
OAG2	17.85	[9.8 - 9.83]
Age2:OAG1	1.09	[-0.5 - -0.54]
Age2:OAG2	-1.83	[-3.0 - -3.02]
Sex:OAG1	2.34	[-2.0 - -2.03]
Q1:OAG1	2.10	[-1.4 - -1.35]
Q4:OAG1	1.75	[-2.9 - -2.94]
Q1:OAG2	-2.62	[-6.3 - -6.28]
Q3:OAG2	2.48	[-0.4 - -0.41]
Q4:OAG2	1.85	[-1.8 - -1.78]
Ed1:OAG1	9.35	[4.7 - 4.72]
Ed2:OAG1	1.62	[-2.7 - -2.71]
Ed3:OAG1	8.08	[-1.0 - -1.03]
Ed2:OAG2	-2.36	[-5.2 - -5.21]
Ed3:OAG2	9.58	[0.2 - 0.16]
ME2:OAG1	7.25	[-0.7 - -0.66]
ME1:OAG2	-11.49	[-28.7 - -28.70]
ME3:OAG2	-2.21	[-5.0 - -4.97]
Oc1:OAG1	-4.95	[-10.5 - -10.54]
Oc2:OAG1	-5.29	[-9.9 - -9.86]
Depre:OAG1	-20.13	[-27.9 - -27.88]
Hta:OAG1	2.99	[-0.3 - -0.31]
Depre:OAG2	-7.26	[-14.8 - -14.82]
Hta:OAG2	3.58	[0.9 - 0.89]

	n	4819
	Ghana	
	Coef	**CI**
(Intercept)	-1.44	[-5.3 - -5.30]
Age2	4.40	[3.9 - 3.87]
Sex	1.56	[-0.1 - -0.13]
Zone	2.97	[1.3 - 1.34]
ME1	3.02	[-0.3 - -0.29]
ME2	2.01	[-1.0 - -1.04]
ME3	3.81	[1.7 - 1.67]
Ed1	0.10	[-2.2 - -2.22]
Ed2	-0.32	[-2.3 - -2.29]
Ed3	0.89	[-2.7 - -2.67]
Q1	0.58	[-2.0 - -2.01]
Q2	0.42	[-2.4 - -2.37]
Q3	-3.86	[-6.6 - -6.57]
Q4	-2.53	[-5.3 - -5.29]
Oc1	-3.00	[-7.4 - -7.42]
Oc2	6.74	[4.8 - 4.75]
Depre	8.90	[4.4 - 4.35]
Hta	-0.66	[-2.3 - -2.33]
OAG1	28.31	[20.4 - 20.41]
OAG2	24.39	[11.0 - 10.95]
Age2:OAG1	-4.04	[-5.1 - -5.13]
Age2:OAG2	-3.47	[-5.4 - -5.41]
Q1:OAG1	5.46	[0.3 - 0.28]
Q2:OAG1	8.34	[3.1 - 3.13]
Q3:OAG1	9.84	[4.9 - 4.89]
Q4:OAG1	4.55	[-1.1 - -1.06]
Q2:OAG2	7.79	[0.5 - 0.50]
Q3:OAG2	10.84	[3.8 - 3.81]
Q4:OAG2	4.76	[-1.9 - -1.93]
Ed1:OAG1	-6.55	[-11.3 - -11.28]
Ed2:OAG1	-3.84	[-8.7 - -8.67]
Ed1:OAG2	-9.59	[-17.1 - -17.14]
ME1:OAG1	5.02	[-1.9 - -1.86]
ME2:OAG1	-6.10	[-12.1 - -12.10]
ME3:OAG2	-2.90	[-10.3 - -10.31]
Oc1:OAG1	-9.48	[-16.1 - -16.11]
Depre:OAG1	-10.88	[-16.6 - -16.60]
Hta:OAG1	2.87	[-0.8 - -0.75]
Depre:OAG2	-9.29	[-16.8 - -16.83]
Hta:OAG2	6.44	[1.4 - 1.39]

	n	10846
	India	
	Coef	**CI**
(Intercept)	7.55	[5.4 - 5.40]
Age2	4.14	[3.9 - 3.85]
Sex	3.64	[2.6 - 2.62]
Zone	1.74	[0.8 - 0.76]
ME1	-0.19	[-1.7 - -1.70]
ME2	1.75	[-2.6 - -2.58]
ME3	1.57	[0.3 - 0.28]
Ed1	-0.33	[-1.5 - -1.52]
Ed2	-2.21	[-3.3 - -3.27]
Ed3	-4.73	[-6.4 - -6.45]
Q1	-0.04	[-1.3 - -1.34]
Q2	-1.46	[-2.8 - -2.75]
Q3	-2.45	[-3.8 - -3.80]
Q4	-4.48	[-5.8 - -5.84]
Oc1	1.09	[0.1 - 0.05]
Oc2	1.72	[0.7 - 0.66]
Depre	13.15	[11.5 - 11.50]
Hta	1.93	[0.8 - 0.77]
OAG1	14.33	[3.2 - 3.25]
OAG2	18.08	[8.9 - 8.94]
Age2:OAG1	-1.43	[-2.5 - -2.54]
Age2:OAG2	-1.72	[-2.8 - -2.80]
Sex:OAG1	1.02	[-2.8 - -2.83]
Sex:OAG2	1.83	[-1.6 - -1.63]
Q1:OAG1	-3.86	[-8.7 - -8.68]
Q2:OAG1	-1.09	[-6.6 - -6.63]
Q3:OAG1	1.39	[-3.1 - -3.09]
Q4:OAG1	-0.16	[-5.2 - -5.21]
Q1:OAG2	-5.94	[-10.6 - -10.56]
Q2:OAG2	-5.60	[-10.5 - -10.54]
Q3:OAG2	-8.23	[-13.4 - -13.38]
Q4:OAG2	-5.70	[-11.0 - -11.03]
Ed1:OAG1	-1.44	[-7.3 - -7.30]
Ed2:OAG1	-2.36	[-6.9 - -6.92]
Ed3:OAG1	-0.83	[-9.1 - -9.08]
Ed1:OAG2	1.56	[-2.2 - -2.18]
Ed2:OAG2	0.71	[-4.0 - -3.99]
Ed3:OAG2	2.57	[-4.0 - -3.99]
ME1:OAG1	-3.92	[-10.6 - -10.62]
ME2:OAG1	-12.44	[-20.0 - -20.05]
ME3:OAG1	-1.92	[-5.9 - -5.87]
ME1:OAG2	-5.64	[-12.3 - -12.26]
ME2:OAG2	-10.64	[-23.2 - -23.18]

ME3:OAG2	-1.30	[-4.8 - -4.83]
Oc1:OAG2	0.18	[-3.7 - -3.74]
Oc2:OAG2	6.67	[2.6 - 2.56]
Oc1:OAG1	-1.85	[-5.7 - -5.72]
Oc2:OAG1	3.18	[-0.3 - -0.27]
Zone:OAG1	0.43	[-5.1 - -5.15]
Zone:OAG2	0.11	[-4.2 - -4.23]
Depre:OAG1	-7.32	[-10.8 - -10.79]
Hta:OAG1	0.56	[-2.5 - -2.48]
Depre:OAG2	-2.01	[-7.1 - -7.09]
Hta:OAG2	0.13	[-3.3 - -3.32]

Table S5.18. Continuation

	n	2451
	Mexico	
	Coef	**CI**
(Intercept)	14.13	[7.6 - 7.59]
Age2	2.47	[1.6 - 1.64]
Sex	4.02	[1.5 - 1.54]
Zone	2.05	[-1.1 - -1.09]
ME1	0.21	[-2.6 - -2.65]
ME2	2.94	[-1.1 - -1.06]
ME3	0.80	[-1.7 - -1.71]
Ed1	-2.09	[-5.4 - -5.43]
Ed2	-4.72	[-7.9 - -7.94]
Ed3	-3.00	[-7.3 - -7.31]
Q1	-2.65	[-7.1 - -7.11]
Q2	-3.65	[-7.7 - -7.70]
Q3	-6.03	[-9.8 - -9.79]
Q4	-6.53	[-10.6 - -10.58]
Oc1	-0.37	[-3.3 - -3.30]
Oc2	1.74	[-1.0 - -1.02]
Depre	8.35	[4.8 - 4.81]
Hta	-1.26	[-3.7 - -3.66]
OAG1	10.68	[3.4 - 3.45]
OAG2	23.52	[9.3 - 9.34]
Age2:OAG2	-3.42	[-5.6 - -5.65]
Sex:OAG1	2.63	[-2.3 - -2.28]
Sex:OAG2	-3.42	[-10.0 - -10.02]
Q1:OAG1	-8.72	[-17.0 - -17.02]
Q2:OAG1	-2.38	[-8.8 - -8.75]
Q3:OAG1	-1.39	[-7.6 - -7.60]
Q4:OAG1	-10.95	[-18.3 - -18.34]
Q1:OAG2	5.44	[-3.3 - -3.29]
Q2:OAG2	13.21	[4.1 - 4.12]
Q3:OAG2	2.53	[-5.8 - -5.79]
Q4:OAG2	16.42	[7.3 - 7.35]
Ed2:OAG1	7.60	[2.0 - 2.00]
Ed3:OAG1	-2.07	[-9.1 - -9.11]
Ed1:OAG2	-12.30	[-18.6 - -18.60]
Ed2:OAG2	-1.91	[-8.9 - -8.89]
Ed3:OAG2	-7.12	[-18.4 - -18.44]
ME2:OAG1	-17.36	[-26.2 - -26.19]
ME2:OAG2	9.43	[0.2 - 0.17]
ME3:OAG2	-1.81	[-9.4 - -9.37]
Oc1:OAG2	-9.11	[-15.5 - -15.50]
Oc2:OAG2	-4.70	[-12.3 - -12.31]
Oc1:OAG1	-2.49	[-8.1 - -8.13]
Oc2:OAG1	1.64	[-3.4 - -3.43]
Zone:OAG1	-6.65	[-11.7 - -11.72]
Zone:OAG2	-2.93	[-10.0 - -10.00]

	n	3603
	Russia	
	Coef	**CI**
(Intercept)	11.76	[3.0 - 2.96]
Age2	4.17	[3.2 - 3.19]
Sex	-1.24	[-4.6 - -4.65]
Zone	5.41	[2.4 - 2.43]
ME1	0.40	[-4.5 - -4.47]
ME2	3.84	[0.2 - 0.15]
ME3	1.29	[-1.3 - -1.27]
Ed1	-9.27	[-14.9 - -14.90]
Ed2	-11.71	[-17.1 - -17.11]
Ed3	-11.27	[-17.6 - -17.57]
Q1	2.61	[-0.6 - -0.61]
Q2	0.91	[-2.1 - -2.08]
Q3	-2.51	[-6.4 - -6.37]
Q4	-1.23	[-5.1 - -5.12]
Oc1	4.29	[-3.6 - -3.62]
Oc2	6.09	[3.9 - 3.86]
Depre	12.08	[8.8 - 8.79]
Hta	2.44	[-1.9 - -1.90]
OAG1	13.49	[6.5 - 6.51]
OAG2	7.36	[-7.1 - -7.07]
Age2:OAG1	-0.95	[-2.4 - -2.37]
Age2:OAG2	-1.28	[-3.3 - -3.26]
Sex:OAG2	2.97	[-1.1 - -1.07]
Q1:OAG1	-3.53	[-8.2 - -8.16]
Q3:OAG1	5.69	[0.6 - 0.62]
Q4:OAG1	3.88	[-1.7 - -1.67]
Q1:OAG2	-10.45	[-15.5 - -15.50]
Q2:OAG2	-7.37	[-11.7 - -11.74]
Ed2:OAG1	-2.25	[-6.4 - -6.42]
Ed1:OAG2	10.54	[2.3 - 2.28]
Ed2:OAG2	11.34	[4.3 - 4.27]
Ed3:OAG2	11.10	[2.9 - 2.89]
ME1:OAG1	9.00	[3.0 - 3.03]
ME2:OAG1	3.27	[-1.8 - -1.85]
ME1:OAG2	9.85	[-1.2 - -1.18]
ME2:OAG2	-11.90	[-17.4 - -17.41]
Oc2:OAG2	2.59	[-2.2 - -2.15]
Zone:OAG1	-2.99	[-7.4 - -7.38]
Zone:OAG2	-4.84	[-9.0 - -8.99]
Hta:OAG1	-2.73	[-8.1 - -8.09]
Hta:OAG2	-1.94	[-8.0 - -8.00]

	n	2596
	South Africa	
	Coef	**CI**
(Intercept)	-0.72	[-11.3 - -11.31]
Age2	4.32	[3.0 - 3.04]
Sex	-0.02	[-5.9 - -5.88]
Zone	-5.78	[-11.3 - -11.29]
ME1	3.47	[-3.9 - -3.87]
ME2	1.94	[-4.8 - -4.75]
ME3	-1.05	[-5.7 - -5.75]
Ed1	-2.94	[-8.0 - -8.05]
Ed2	-0.91	[-6.6 - -6.59]
Ed3	0.55	[-5.8 - -5.76]
Q1	1.35	[-4.8 - -4.80]
Q2	-1.38	[-10.3 - -10.26]
Q3	-3.29	[-8.1 - -8.05]
Q4	-5.01	[-11.3 - -11.32]
Oc1	6.53	[0.3 - 0.30]
Oc2	6.74	[1.7 - 1.75]
Depre	15.40	[8.1 - 8.14]
Hta	3.78	[-1.1 - -1.09]
OAG1	37.30	[26.3 - 26.26]
OAG2	15.43	[6.3 - 6.35]
Age2:OAG1	-5.55	[-7.5 - -7.46]
Sex:OAG1	-5.26	[-12.4 - -12.39]
Q2:OAG1	3.69	[-5.8 - -5.80]
Q4:OAG1	6.99	[-2.1 - -2.13]
Q1:OAG2	-10.22	[-18.4 - -18.42]
Q4:OAG2	-3.82	[-11.9 - -11.94]
Ed1:OAG2	6.92	[-0.5 - -0.50]
Ed2:OAG2	-4.14	[-11.9 - -11.94]
ME1:OAG1	-14.72	[-22.9 - -22.90]
ME3:OAG1	8.25	[1.7 - 1.70]
ME1:OAG2	6.08	[-2.3 - -2.34]
Oc2:OAG2	-3.21	[-9.1 - -9.13]
Zone:OAG1	9.90	[2.2 - 2.21]
Zone:OAG2	5.52	[-1.5 - -1.46]
Hta:OAG2	-5.40	[-13.7 - -13.75]

- *Age2: age*
- *ME: marital status (1: never married; 2:separated or divorced; 3: widowed; reference: married)*
- *Ed: education (1: primary completed; 2 secondary completed; 3 college or university completed; reference: less than primary)*
- *Q: quintil of income*
- *Oc: occupation (1: never worked/ homemaker; 2: not working; reference: working)*
- *Depre: depression*
- *OA: osteoarthrosis*
- *HtaG1: hypertension without treatment*
- *HtaG2: hypertension with treatment*

Depre:OAG1	5.20	[-1.4 - -1.41]
Depre:OAG2	3.52	[-2.2 - -2.20]
Hta:OAG2	9.11	[3.4 - 3.41]

Figure S5.21. Histograms of observed health gain attributable to healthcare interventions addressed to osteoarthritis by country. SAGE study. wave 1.

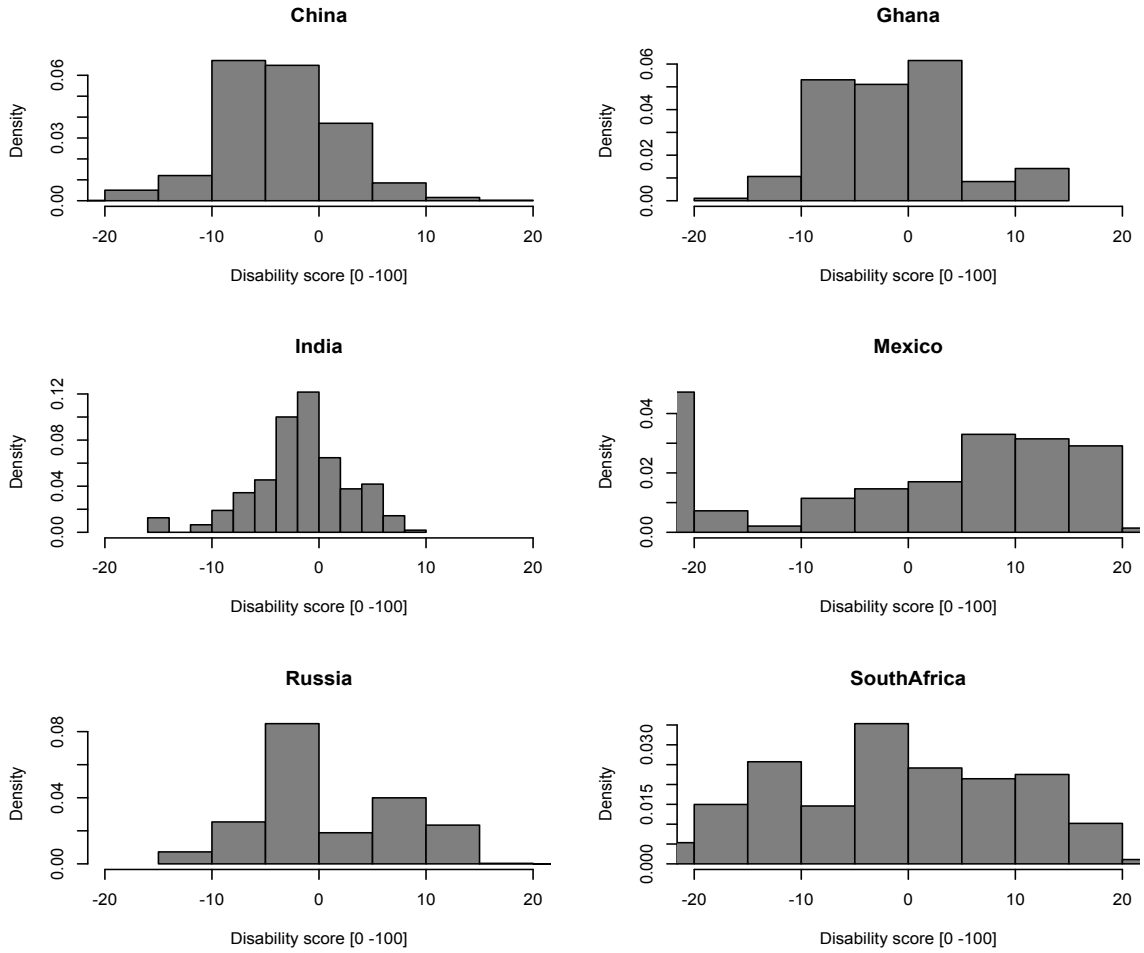
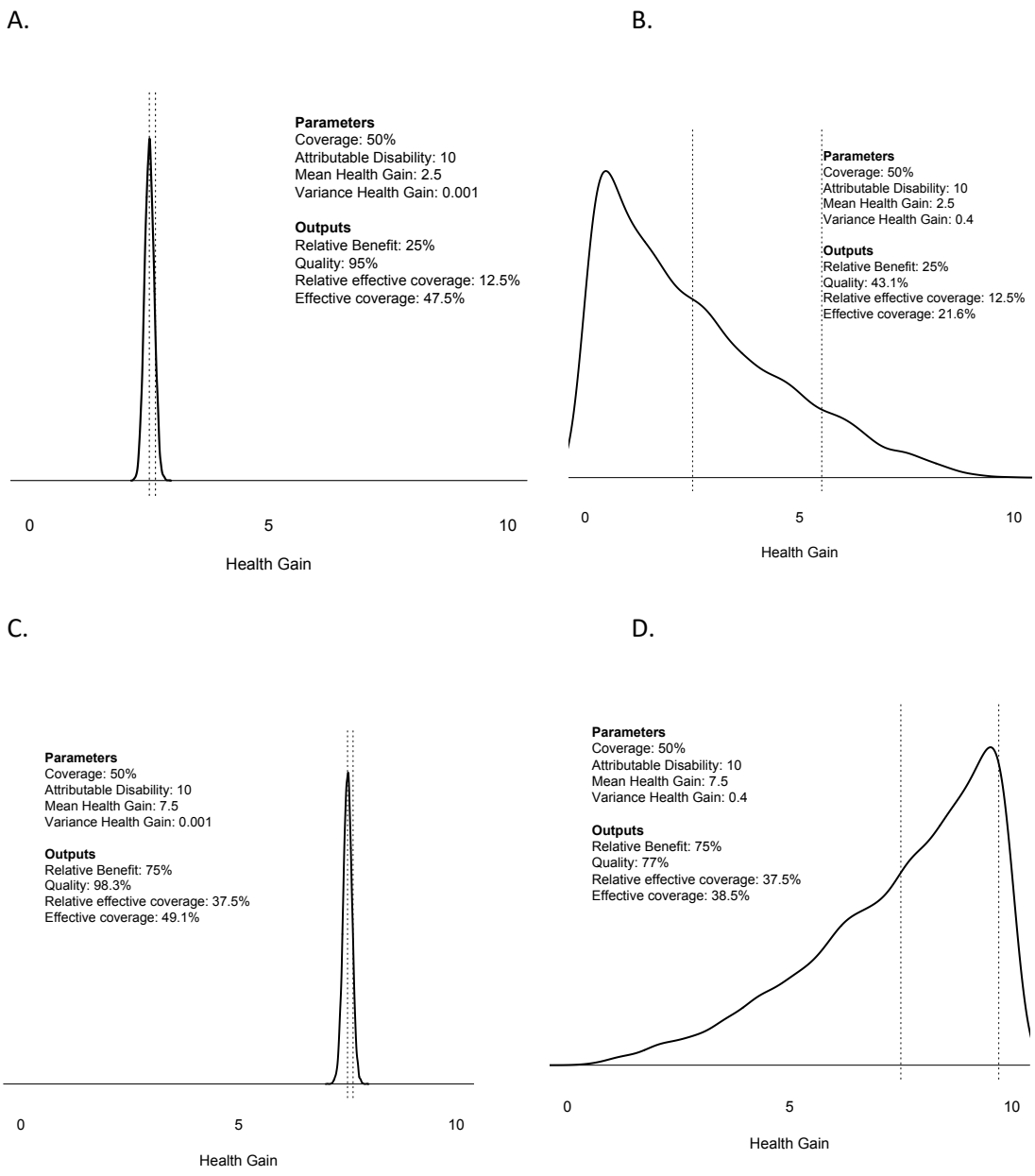


Table S5.19. Results of effective coverage of the healthcare for osteoarthrosis, by country, assuming effective coverage as an individual attribute. SAGE study, wave 1.

	n=13311		n=4819		n=10846	
	China		Ghana		India	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	34.0	[-136.8 - 204.7]	114.9	[-61.7 - 291.5]	-748.2	[-1832.7 - 336.2]
Quality (%)	-152.9	[-196 - -109.7]	-38.9	[-78.3 - 0.5]	-43.2	[-59.7 - -26.7]
relative - effective coverage (%)	18.2	[-73.4 - 109.7]	28.7	[-14.3 - 71.7]	-323.3	[-792.3 - 145.7]
absolute - effective coverage (%)	1.5	[-6.2 - 9.3]	2.8	[-1.4 - 7]	-28.2	[-69.2 - 12.7]
Effective coverage (%)	-81.8	[-108.8 - -54.7]	-9.7	[-19.8 - 0.4]	-18.7	[-26.1 - -11.2]
RR HG-average	1.51	[0.42 - Inf]	Inf	[0.62 - Inf]	0.12	[0.05 - Inf]
RR HG-max	2.08	[0.95 - Inf]	0.95	[0.42 - Inf]	0.16	[0.04 - Inf]

	n=2451		n=3603		n=3056	
	Mexico		Russian Federation		South Africa	
	Estimate	CI	Estimate	CI	Estimate	CI
Relative benefit (%)	391.6	[106.4 - 676.7]	18.0	[-44.7 - 80.7]	106.7	[18.8 - 194.6]
Quality (%)	-7.7	[-69.8 - 54.3]	8.6	[-18.8 - 36.1]	-36.1	[-70.4 - -1.8]
relative - effective coverage (%)	186.3	[28.9 - 343.7]	13.0	[-32.9 - 58.9]	64.4	[6.1 - 122.7]
absolute - effective coverage (%)	11.6	[0.3 - 23]	1.8	[-4.7 - 8.2]	6.4	[0.2 - 12.5]
Effective coverage (%)	-3.7	[-33.5 - 26.2]	6.2	[-14.2 - 26.7]	-21.8	[-43.8 - 0.2]
RR HG-average	Inf	[Inf - Inf]	1.22	[0.69 - 5.18]	Inf	[1.23 - Inf]
RR HG-max	0.29	[0.11 - Inf]	5.02	[0.33 - Inf]	1.20	[0.54 - Inf]

Figure S5.22. Hypothetical health gain distributions and results for relative benefit, quality, relative effective, and effective coverage.*



*Health Gain were simulated restricting values between 0 and 10, using a Beta distribution. Means and variance of the distribution are shown in each panel. Raw coverage is assumed 50% and the attributable disability at the individual level, equal to 10.