Development of a Toolkit for the Assessment and Monitoring of Disabilities due to Neglected Tropical Diseases

Getting a first impression of its validity and usefulness in Brazil

Student:
Heleen Kuiper (2535496)
h2.kuiper@student.vu.nl
Master International Public Health

Placement:
Netherlands Leprosy Relief**/** Neglected Tropical Disease NGDO Network

Supervisors:
J.W. Brandsma, PhD
W.H. Van Brakel, PhD

September 2015

This report focuses on aspects of validity, usefulness and relevance of two tools: 1. the WHO Disability Assessment Schedule (WHODAS) 2.0 12-items, assessing functional limitations and 2. the Participation-scale (Short), assessing participation restrictions. Both validated among persons with Chagas disease, leishmaniasis, leprosy or schistosomiasis in Brazil.

ABSTRACT

Background Neglected tropical diseases (NTDs) are a group of tropical diseases affecting over a billion people worldwide. People affected often live in poverty and have to deal with the disabling consequences of many of these diseases. The similarities in disabling consequences, geographical spread, prevention methods and treatment of NTDs, ask for an integrated control. However, international comparable data about NTD-related morbidity and disability to improve such integrated actions are lacking.

Methods A mixed-methods research design was used to develop a generic, cross-NTD toolkit aimed at assessing and monitoring NTD-related morbidity and disability. The most important phases were an international Delphi study and an initial validation. The Delphi study among NTD and disability experts consisted of three iterative rounds, aimed at investigating priority areas and needed tools in the NTD field. The initial validation of the prototype toolkit took place in the Ceará state of Brazil, conducting interview administered versions of the tools that were ranked by the Delphi study experts. The 34 participants included were affected by at least one of the NTDs occurring in this state.

Results The Delphi study resulted in a prototype toolkit that covered the domains from the International Classification of Functioning, Disability and Health (ICF). The tools that were ranked highest within a domain were included in the toolkit. This report focuses on aspects of validity, usefulness and relevance of two of these tools: the WHO Disability Assessment Schedule (WHODAS) 2.0 12-items, included for the assessment of functional limitations and the Participations-scale Short, included for the assessment of participation restrictions. Results from the initial validations suggest those tools will be useful among persons with Chagas disease, leishmaniasis, leprosy or schistosomiasis.

Discussion and conclusion Assessment and monitoring of functional limitations and participations restrictions due to Chagas disease, leishmaniasis, leprosy or schistosomiasis can be done with the WHODAS 2.0 12-items and the P-scale Short. It is recommended to further validate these (and other) tools among other NTDs in differing environments and cultures.

ABBREVIATIONS

CHIEF	The Craig Hospital Inventory of Environmental Factors
EMIC	Explanatory Model Interview Catalogue
LF	Lymphatic Filariasis
MMD	Morbidity Management & Disability
NTD	Neglected Tropical Disease
NNN	NTD Non-Governmental Development Organization Network
P-scale	Participation scale
ICF	International Classification of Functioning, Disability and Health
QoL	Quality of life
SRQ	Self-reported questionnaire
SALSA	Screening of activity limitations and safety awareness
WHO	World Health Organization
WHODAS	WHO Disability Assessment Schedule

TABLE OF CONTENTS

1. INTRODUCTION

The global public health community increasingly recognizes the morbidity and disability resulting from neglected tropical diseases (NTDs). NTDs are a group of 18 tropical diseases, endemic in 149 countries (WHO, 2012). Globally, over 1.4 billion people are affected by NTDs (WHO, 2015) and they often live in poverty (Hotez et al., 2006; Barry, 2014) and social exclusion (WHO, 2013 A). Poverty and social exclusion can be both a cause and a result from an NTD; NTDs could/can affect people in their daily life, for example due to participation problems (Hotez, 2006; Weiss, 2008; Zeeuw, 2014) or disfigurement (Hotez, 2008). Disability could be defined as *"…an umbrella term for impairments, activity limitations and participation restrictions. It denotes the negative aspects of the interaction between an individual (with a health condition) and that individual's contextual factors (environmental and personal factors)."* (WHO, 2013 B).

Disability as described above has a variety of expressions among NTDs. Skin lesions and disfigurement, for example, may result from ulcers and scarring due to leprosy, Buruli ulcer and cutaneous leishmaniasis (WHO, 2015). Visual impairment and even blindness can be a consequence of trachoma and is, in many cases, preventable and curable when detected early (Saxena et al., 2015). Other NTDs that can cause visual impairments are onchocerciasis and leprosy (WHO, 2015). Neurologic manifestations due to NTDs are well-described for Chagas disease, cysticercosis, leprosy, rabies and schistosomiasis, however are expected to appear in other NTDs as well (Berkowitz, 2015). A mobility affecting example is severe lymphedema in lower limbs due to podoconiosis and lymphatic filariasis (LF) (WHO, 2015). As result of the disabling consequences and fear of contagiousness, NTDs may cause stigma (Steinstra et al., 2002; Weiss et al., Van Brakel et al., 2012) and mental health problems (Litt et al., 2012).

The disabling consequences of NTDs are similar for some and so are their geographical spread and methods for detection, prevention, treatment, rehabilitation and surveillance. Preventive single-dose medicines for example could be used to prevent LF, onchocerciasis, schistosomiasis, soil-transmitted helminthiases (STH) as well as trachoma. Vector control, safe water, sanitation and hygiene and capacity strengthening can be integrated to control NTDs worldwide (WHO, 2012; WHO, 2013A). Several such integrated programmes have been set up and have reported improved delivery in resource-poor countries (Grépin & Reich, 2008), better cost-effectiveness (Brady et al., 2006; Hotez) and earlier identification of individuals and communities at high risk for disability and morbidity (Lammie et al., 2006).

And yet, internationally comparable data about NTD-related morbidity and disability to improve such integrated actions are lacking (Mont, 2007; Van Brakel & Officer, 2008). This shortcoming is confirmed by participants in the recent (February, 2015) NTD Cross-cutting Issues Workshop in Utrecht, The Netherlands and the Neglected Tropical Disease Non-Governmental Development Organization Network (NNN). These experts identified the need for development of a generic cross-NTD toolkit to assess and monitor NTD-related morbidity and disability, as one of the key priorities. Improved availability of reliable data, in addition to the possibilities for integrated control, is of great importance for funding, planning, implementation, monitoring and evaluation of existing and new appropriate interventions and policies where needed (Mont, 2007). In summary, a reliable generic cross-NTD toolkit is expected to provide essential information for the identification of priority areas to reduce suffering in people with NTD-related disabilities. Here, Morbidity Management and Disability (MMD) is an often used phrase when referring to management of the consequences of NTDs (WHO, 2013 C).

In recent years, disability-experts increasingly recognized the ‘International Classification of Functioning, Disability and Health’ (ICF) as a useful classification on the basis of which assessment instruments can be structured (Van Brakel et al., 2006; Mont, 2007; Van Brakel & Officer, 2008). This classification supports a worldwide understanding of the concept ‘disability’. The ICF is a classification system that can be used to develop globally standardized assessment tools able to measure overall and domain-related health and disability in individuals and populations (Eide & Loeb, 2005). The ICF comprises six interacting components: health condition, body functions and structures, activity, participation, environmental and personal factors (WHO, 2002). One motive for the use of the ICF is that environmental and personal factors are of special importance in developing a generic toolkit that aims to be workable worldwide within and across different contexts. Such contextual factors, for example, differ in demographics, cultures, governmental systems, race, gender, age and educational level. Contextual factors determine ‘outcome’ and need to be known in order to interpret and compare data. Therefore, it is highly important to consider these contextual factors in the assessment of disability.

Another important strength of the ICF is that the development of the ICF was based on extensive cross-cultural research (Üstün, 2001). This is important because concepts and items used in assessment tools may be understood differently between cultures and languages (Herdmann et al., 1998). Since the importance of international comparable data on disabilities (Mont, 2007), Stevelink & Van Brakel (2013) argue that it is crucial to test the cross-cultural applicability of concepts and items used within tools. For this purpose, they developed a checklist for the assessment of reporting on cross-cultural validation based on the work of Herdmann (1998) (Appendix 1).

Given the international recognition of the ICF, this classification system is used in accomplishing the overall aim of this study: *Developing a generic cross-neglected tropical disease (NTD) toolkit, aimed to assess and monitor overall disability and domain-specific disability due to NTDs and by doing this to identify priority areas for morbidity management and disability (MMD) services*. In order to meet this objective, three research phases can be distinguished: (1) a Delphi study to investigate and finally present a prototype toolkit; (2) a systematic literature review to test cultural equivalence of prioritized tools; and (3) a pilot and initial validation of the toolkit in Brazil, to describe aspects of validity, usefulness and relevance of two of its tools.

2. BACKGROUND

2.1. Theoretical Background

2.1.1. Neglected tropical diseases (NTDs)

NTDs are a group of 17 tropical diseases, prioritized by the World Health Organization (WHO) and endemic in many countries (WHO, 2012). For this study, the list of 17 is extended with Podoconiosis, another NTD with disabling consequences, nonetheless, 'double neglected' at international level because this NTD is not prioritized by the WHO (Davey, 2010). This group of 18 diseases, as can be seen in Table 1, are neglected in terms of attention and interventions by governments, pharmaceutical industries, research funds, etc. (Barry, 2014).

Table 1. Overview of Neglected Tropical Diseases

The 17 prioritized NTDs by the WHO and Podoconiosis, divided into causative pathogens (WHO, 2015).

Protozoa	Bacteria	Helminthes	Viruses	Other
Chagas disease Human African trypanosomiasis (sleeping sickness) Leishmaniasis	Buruli ulcer Leprosy (Hansen disease) Trachoma Yaws	Cysticercosis/ Taeniasis Dracunculiasis (guinea-worm disease) Echinococcosis Foodborne trematodiases Lymphatic Filariasis Onchocerciasis (river blindness) Schistosomiasis Soil-transmitted helminthiases	Dengue and Chikungunyu Rabies	Podoconiosis

According to the WHO (2012), NTDs are endemic in 149 countries with differing populations and environments. The most vulnerable populations can be found in 10 global hotspots: Brazil & the Amazon Region, Gran Chaco, Mesoamerica &Texas, Nigeria, Democratic republic of Congo

(DRC) & adjoining nations, Chad, Niger, Mali, & adjoining Sahelian Areas, Indonesia & Papua New Guinea, India & South Asia, China, Middle East (Hotez, 2014) (figure 1).

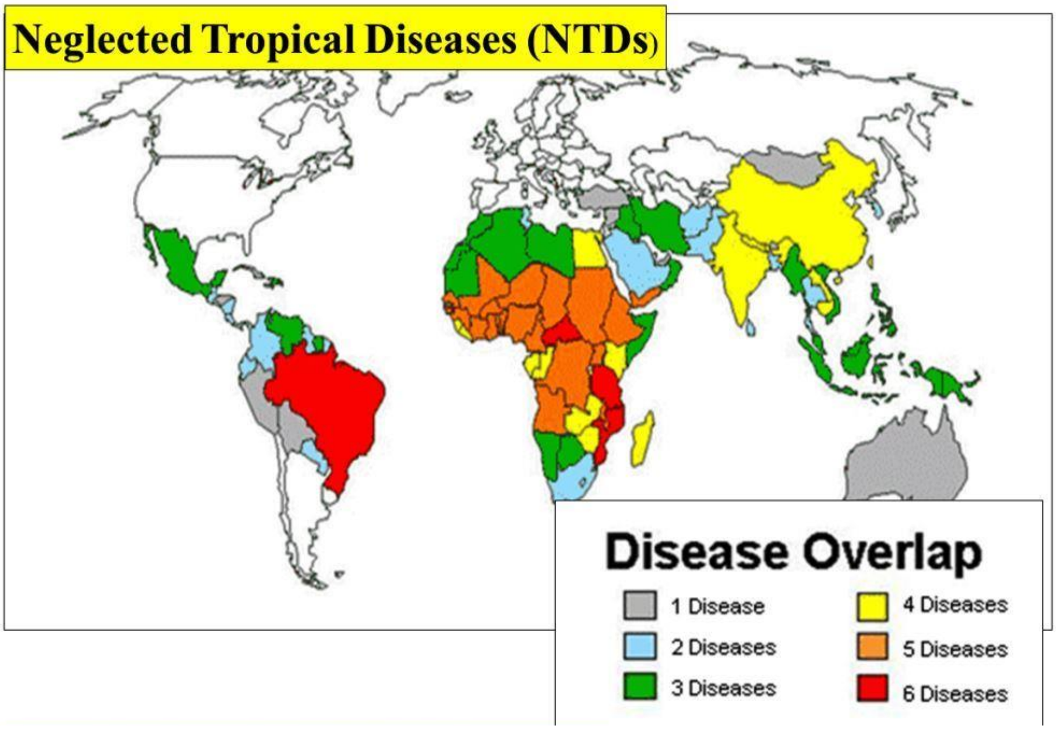


Figure 1. Global burden of NTDs.

2.1.2. Disability

Disability can by defined by a variety of models or approaches. These definitions are essential in the assessment of disability. Assessment of disability could, for example, be focused on disabilities in individual bodies; this focus is described by the medical model (Davis, 2006). The medical model suggests that disability is the inability to function within an individual. A broader and more recent approach is the social model which does not describes the inability to function on its own, but which also includes the interaction with the physical, cultural and political environment (Mont, 2007). Despite the shift from a medical model to a social model, both models still suggest dichotomously; either purely medical or purely social. Therefore, there is a need for more comprehensive approaches that combine disability and its overall impact. Such a combination is made in the approach of the 'International Classification of Functioning, Disability and Health' (ICF) (WHO, 2011 A). The ICF defines disability as *"…an umbrella term for impairments, activity limitations and participation restrictions. It denotes the negative aspects of the interaction between an individual (with a health condition) and that individual's contextual factors (environmental and personal factors)."* (WHO, 2013 A).

2.1.3. Morbidity

Morbidity is defined as *"Any departure, subjective or objective, from a state of physiological or psychological well-being"* (Last, 2001). This definition covers any health statuses on biological, individual and social levels and it addresses medical problems as well as functional disabilities and subjective health. Beside this definition, morbidity is also used as a unit that says something about the number of individuals that have a disease, the illnesses that these individuals perceive and the duration of the disease (Last, 2001).

2.1.4. Morbidity Management & Disability (MMD)

In 2013, the World Health Organization (WHO) published a report about morbidity management and disability prevention (MMDP) in LF (WHO, 2013 C). In the current study the term MMDP as described in the WHO report is adjusted to a cross-NTD approach instead of merely LF focused. Additionally, because of ethical considerations and because persons with disabilities may be in need for inclusive interventions, the "P" of prevention is excluded. Therefore the adapted term 'morbidity management and disability' (MMD) is used for further understanding.

The overall goal of MMD is to reduce suffering in people with NTD-related disabilities and to improve quality of life (QoL) by access to recommended basic care for everyone suffering from the disabling consequences of NTDs. The minimum package of care that is recommended for LF might be applicable for all NTDs. If so, such a cross-NTD package recommends (WHO, 2013 C):

- treating acute NTD-related symptoms;
- preventing debilitating and painful episodes of acute NTD-related symptoms and progression of other symptoms;
- providing access to surgery of NTD-related symptoms/ consequences;
- providing antifilarial medicines by mass drug administration or individual treatment.

Lifelong access to continuing care for NTDs seems important to manage the disease and to prevent progression. According to this and to enhance sustainability, it is also recommended to include MMD in the primary health care systems (Molyneux, 2012).

Planning of MMD in a national programme is done by three guiding principles and policies: (1) 'access', referring to the accessibility of basic care for all persons with (acute) NTD-related symptoms; (2) 'flexibility', referring to allowance of flexible approaches in prevention and reduction of suffering from NTD-related symptoms; (3) 'integration', referring to integration of activities into other disease-control programmes. These principles can be reached by a strategic plan that consists of three steps (Figure 2) (WHO, 2013 C). Appendix 2 gives a description of these steps.

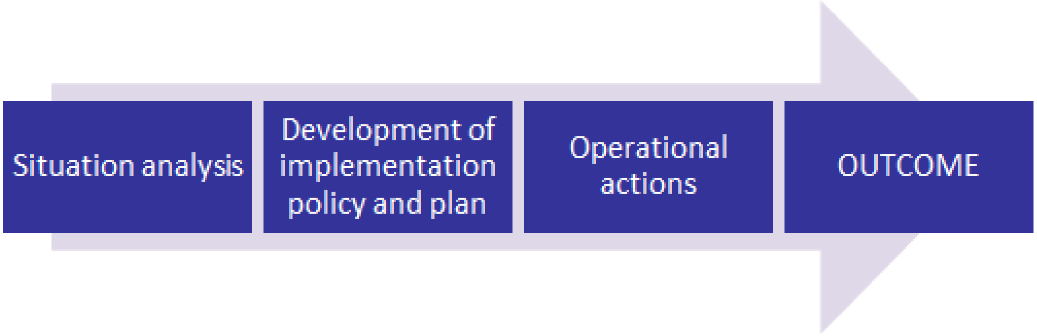


Figure 2. Strategic plan for effective morbidity management and disability (MMD) (WHO, 2013 C).

2.1.5. International Classification of Functioning, Disability and Health (ICF)

The earlier mentioned 'International Classification of Functioning, Disability and Health' (ICF) framework gives insight into three perspectives of health: the biological perspectives, the individual perspectives and the social perspectives. In Figure 3 it can be seen that the functioning of an individual reflects the interaction between health condition and contextual factors; environmental and personal.

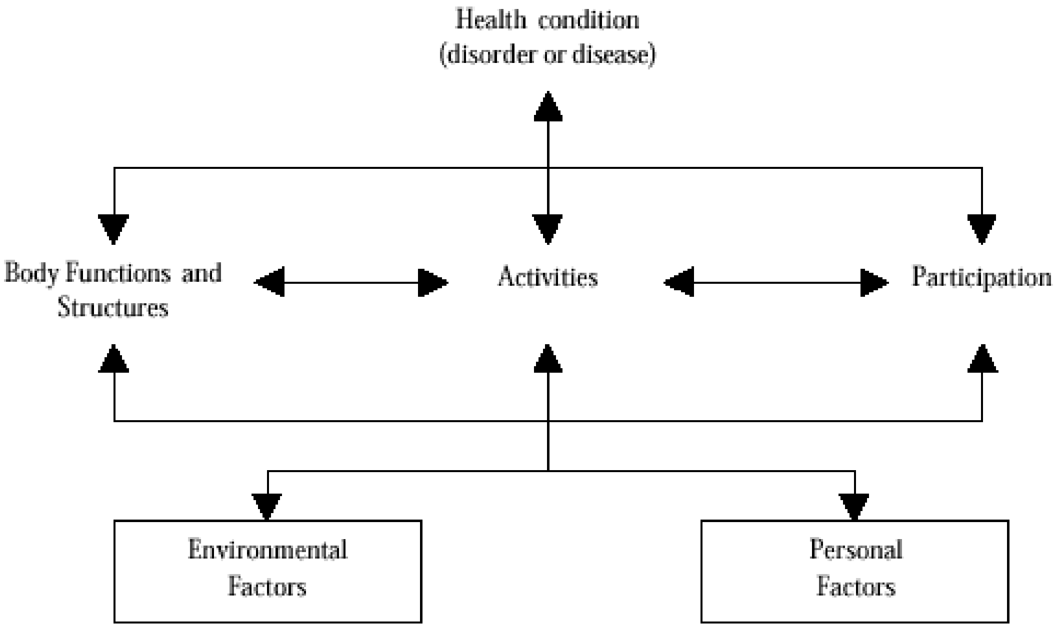


Figure 3: The ICF framework as developed by the WHO in 2001 (WHO, 2002).

The meaning of 'body functions' in this framework covers all physical and psychological functions of body systems while 'body structures' focuses more on the anatomy of the body. Functioning in these levels is related to the ability to execute tasks or actions called 'activities'.

Related to both of the latter functioning levels is ‘participation’ which means involvement in a life situation. In the same way, problems with functioning could be described on different levels: problems with body functions and structures are described as ‘impairments’, problems with activities as ‘functional limitations’ and problems with participation as ‘participation restrictions’. These different levels of problems are included in the definition of ‘disability’ as described before. Disability in individuals as well as in populations therefore could be holistically monitored, assessed and classified by applying the ICF in the development of tools. Hence, data gathered by application of the ICF can support planning of health services and interventions for managing morbidity and preventing disability (WHO, 2013 B).

2.2. Contextual Background

2.2.1. Conceptual Framework

Justifying the development of the Framework

The ICF is used as a guideline to compose a holistic toolkit for worldwide assessment and monitoring of domain-specific and overall morbidity and disability due to NTDs and therefore included in this preparatory development level. The next level, the validation level, comprises testing aspects of validity, usefulness and relevance (and cultural equivalence) of the developed toolkit. This level is important for the cross-NTD character of the toolkit and thus the worldwide use. After this level, action is needed; the first step of the strategic plan for organization of effective MMD services is integrated because it provides a clear example of actions that can be taken to analyse the current situation in an area (WHO, 2013 B).

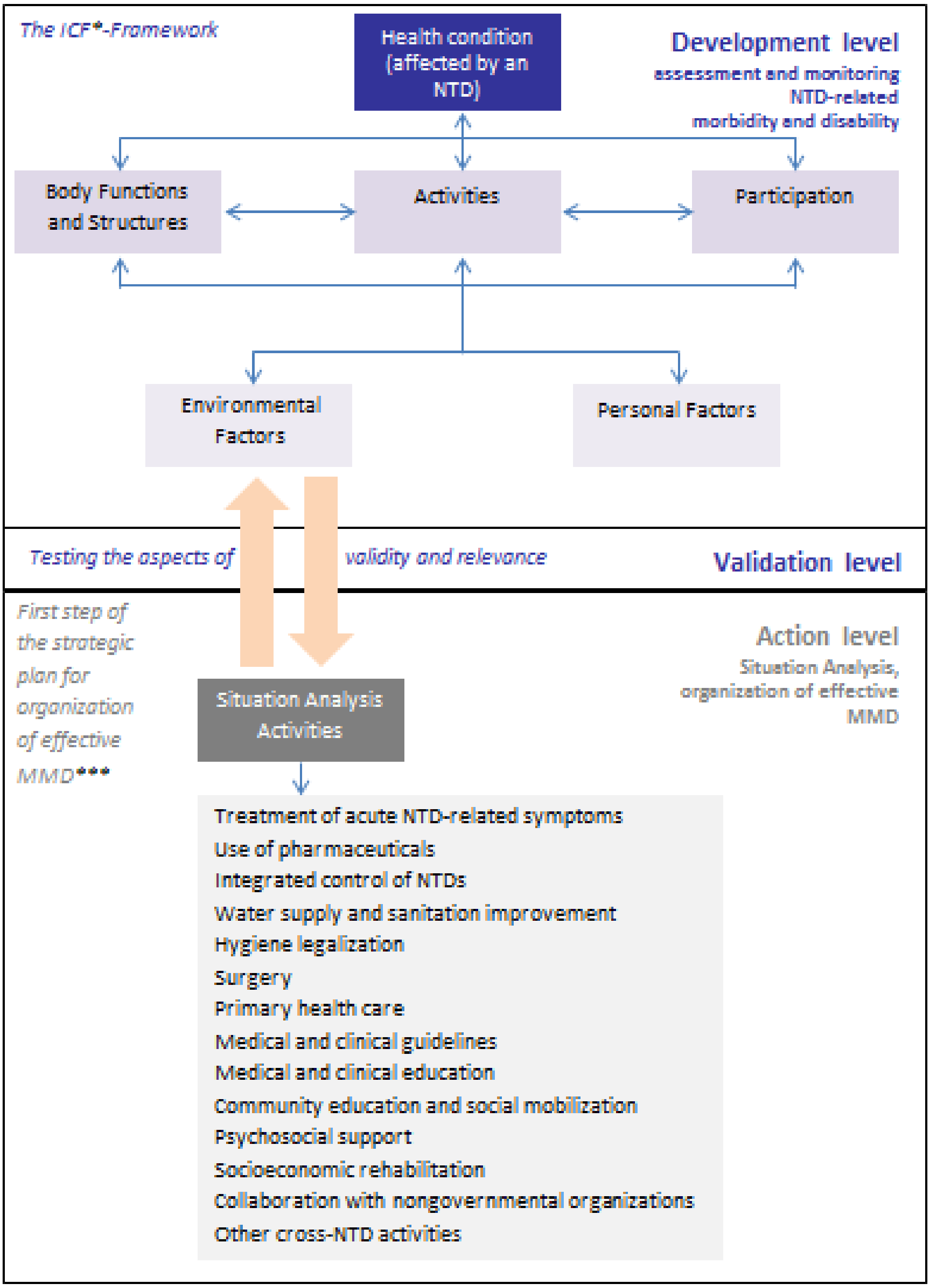


Figure 4. Conceptual framework
The first block displays the ICF on which the development of holistic, generic cross-NTD toolkit is based. The ICF-domain of environmental factors is linked with the third block, in which the first MMD-step is displayed. In between the development level and the action level there is the validation level in which cross-cultural equivalence is tested, before taking action.
* ICF = International Classification of Functioning, Disability and Health (WHO, 2013 B), ** NTD = Neglected tropical disease, *** MMD = Morbidity management and disability (WHO, 2013 C).

Guiding through the framework; understanding the relations

Within the development level of the conceptual framework the concept 'health condition' is the starting point for assessing NTD-related morbidity and disability. In this context health condition may refer to the 18 NTDs included in this study. The morbidity aspect is part of 'body functions and structures'. Accordingly, the disability aspect is determined by the interaction between the health condition, functioning and contextual factors. Functioning of an individual depends on the remaining concepts in this ICF framework: 'body functions and structures', 'activities' and 'participation' (WHO, 2013 B). Gathering data by means of ICF based tools could support the situation analysis as is described in the action level of the conceptual framework. However, these tools, which are meant to be used worldwide, first have to be tested for their validity, usefulness and relevance in differing cultures. After the validation level comes the action level which is not only supported by the development level as described above but parts of it also interact with this level. The situation, as analysed in the action level, could also affect the environment of individuals with NTD-related morbidity and disability.

2.2.2. Research Objectives and Research Questions

Main Objective

Developing a generic cross-neglected tropical disease (NTD) toolkit, aimed to assess and monitor overall disability and domain-specific disability due to NTDs and by doing this to identify priority areas for morbidity management and disability (MMD) services.

Main question: *"What should the content be of a generic cross-NTD toolkit aimed at assessing and monitoring aspects of NTD-related morbidity and disability and identifying priority areas for MMD?"*

Specific Objectives

1. Investigating priority areas and (currently used) tools aimed to assess and monitor NTD-related morbidity and disability, by a Delphi study that asks for the perspectives from experts in the field of NTDs and/ or morbidity and disabilities and/ or monitor and assessment tools.

> Sub-question 1: *"What are the priority areas and (currently used) tools aimed to assess and monitor NTD-related morbidity and disability according to experts in the field of NTDs and/ or morbidity and disabilities and/ or monitor and assessment tools?"*

2. Presenting a prototype of the generic cross-NTD toolkit and accordingly prioritizing tools for the initial validation by asking experts to rank tools within a Delphi study.

> Sub-question 2: *"What should be included in the toolkit according to experts and which of these tools will be prioritized for initial validation?"*

3. Collecting data concerning cross-cultural validation of the tools that have potential to be included in the generic cross-NTD toolkit by means of a systematic literature review.

> Sub-question 3: *"In which countries, in what languages and for what conditions are the tools that have the potential to be included in the generic cross-NTD toolkit already being validated?"*

4. Collecting (cognitive)data among around 50 persons with NTD-related morbidity and disability in Brazil by an initial validation of the toolkit and asking cognitive questions after each tool and the toolkit as a whole to describe aspects of validity, usefulness and relevance of two of the tools.

> Sub-question 4: "*How valid and relevant are the two prioritized tools according to the target group participating in an initial validation?"*

3. METHODS

3.1. Introduction to the study process

Given the complexity of the research, a mixed research design was used. Table 2 summarizes the methods which will be further explained in detail.

Table 2. Study questions, phases and methods (reading-guide)
Answering the four sub-questions of this study is divided over three phases in which differing methods are used. Not all methods and their results are included in this report.

Research-question/ research phase	Phase	Method used
Sub-question 1: What are the priority areas and (currently used) tools aimed to assess and monitor NTD-related morbidity and disability according to experts in the field of NTDs and/ or morbidity and disabilities and/ or monitoring and assessment tools?	Phase 1	**Exploratory literature study**, relevant tools (not elaborated on in this report). **Delphi study**, 1st and 2nd e-survey.
Sub-question 2: What tools should be included in the toolkit according to experts and which of these tools will be prioritized for initial validation?	Phase 1	**Delphi study**, 3rd e-survey.
Sub-question 3: In which countries, in what languages and for what conditions are the tools that have the potential to be included in the generic cross-NTD toolkit, already being validated?	Phase 2 (Background study)	**Systematic literature review** cross-cultural validation tools. (methods and results in appendices)
Sub-question 4: How valid, relevant and useful are the two prioritized tools according to the target group participating in an initial validation?	Phase 3	**Pilot and initial validation**
Main question: What should the content be of a generic cross-NTD toolkit, aimed at assessing and monitoring aspects of NTD-related morbidity and disability and identifying priority areas for MMD?	Phase 1, 2 and 3	*All of the above*

The research started with a literature search, finding information about possible relevant existing tools. This literature search is not further elaborated on in this report because it was conducted by someone else. Therefore this methods section starts with a description of the Delphi study for which three e-surveys were conducted to answer the first and second sub-questions. Next, the second phase aimed at answering the third sub-question by a systematic literature review to gather information about the cross-cultural

equivalence of the potential tools (Appendix 3). The third phase is the initial validation of two of the toolkit tools: the WHO Disability Assessment Schedule (WHODAS) 2.0, 12-items and the Participation-scale Short (P-scale Short).

3.2. Phase 1: Delphi study

The Delphi study is a qualitative method to build consensus among experts. Such consensus is reached by using a series of iterative questionnaires submitted to a regulated sample of experts who remain anonymous to each other (Hsu & Sandford, 2007). Considering the expected homogeneous nature of Delphi study groups, a sample size of ten to fifteen is usually adequate (Cross, 2014).

3.2.1. Purpose and organization

A Delphi study was conducted to collect and prioritize key-informants' ideas about which aspects of NTD-related morbidity and disability should be assessed and monitored and which tools could be used to do this. The Delphi study involved three rounds (Figure 5). Each round consisted of one e-survey. We realized the e-surveys using a free-account version on "www.esurveyspro.com" and introduced them by email providing a link to the survey. We co-ordinated the Delphi study process by assessing consensus and providing 'feedback' within each following e-survey.

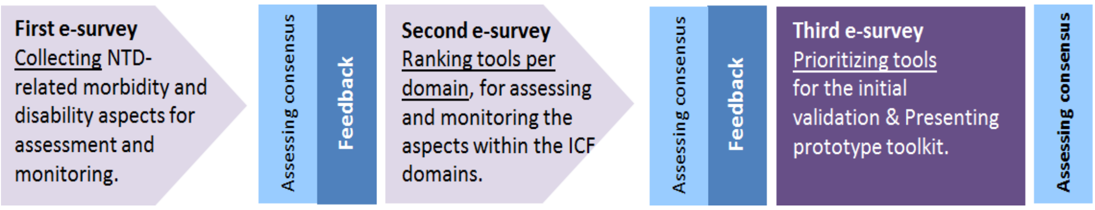


Figure 5. Flow chart of the Delphi study

The flowchart displays the main objectives within the three e-surveys. In this figure, the iterative character of the Delphi study can be easily observed; starting broad in the first e-survey and ending more specific in the last e-survey. After assessing consensus after each e-survey, feedback is provided to the experts within the next e-survey.

3.2.2. Participants

Thirty key-informants were recruited from among the participation list of the NTD Cross-cutting Issues Workshop in Utrecht, Netherlands (2015).

Inclusion criteria

To be included, key-informants had to:

- work closely with NTDs and/or with morbidity and disability and/or have experience with assessment tools in NTD field and/ or be affected with a NTD,
- be involved in management of several NTDs,
- represent a wide range of relevant NTD countries.

Key-informants were invited, introduced to the study and asked for email addresses of additional key-informants in their organization or network. Key-informants that participated in the first round were automatically invited to complete the e-survey for the following round(s). Their characteristics, such as the country in which they work and the NTD(s) they work on, were noticed.

3.2.3. First e-survey

The initial invitation, the research description and the first e-survey were pre-tested among three NTD-experts of whom one also had experience conducting a Delphi study. Comments on the pre-tested e-survey were taken into account to complete the first e-survey (Appendix 4). The first round aimed at obtaining a general overview of NTD-related morbidity and disability aspects which are considered important for assessment and monitoring. Therefore, topics discussed in this e-survey included ‘NTD(s) you are working with’ (Question 1 and 2), ‘importance of ICF domains’ (Question 3 and 5) and ‘NTD-specific and -generic assessment tools’ (Question 4 and 6). Field of expertise was ascertained by listing all eighteen NTDs asking participants to select one or more as their field of expertise. ‘Other’ was also an option which could be explained in an open field. Questions about the six ICF components were also provided as multiple-choice. Open-ended questions were used to collect information about assessment tools for an NTD in which the expert had expertise. An additional question (Question 7) about interventions to facilitate solutions for MMD was included to collect ideas for possible inclusion of a guide for the organization of effective MMD services. At the end of the e-survey, participants were asked to send a copy of the forms of the tools they used.

3.2.4. Second e-survey

The second round of the Delphi study aimed at selecting the most appropriate tools to be included in the generic cross-NTD toolkit. Questions in this e-survey (Appendix 5) were developed on the basis of results from the first e-survey and from the exploratory literature search (phase 1). The most important tools were listed per ICF domain. In order to accomplish a somewhat homogenous level of knowledge about the tools included, questionnaire forms of these tools were attached to this second e-survey. Participants were asked to select or rank tool(s) that in their opinion had the most potential for assessing and monitoring aspects within the ICF-domains. Ranking tools within these domains was achieved by assessing a score according to the perceived priority of the tool. Tools ranked first, were given three points, tools ranked second two, tools ranked third one and tools that were not ranked at all received zero points. Finally, participants were asked to indicate the need for NTD-specific tools.

3.2.5. Third e-Survey

The third round presented the experts with a prototype toolkit, developed by collating answers from the previous two e-surveys (Appendix 6). This prototype toolkit consisted of at least one tool per ICF-domain. Participants were asked to prioritize at least four tools that should be included during the initial validation. Because of the limited timeframe it was impossible to test all the tools. Prioritization was obtained by asking the experts to rank the tools followed by scoring the outcomes. Here, tools ranked first were given four points, tools ranked second three, tools ranked third two and tools ranked fourth one.

3.3. Phase 3: Pilot and initial validation

3.3.1. Purpose and organization

The third phase of this study aimed to collect quantitative and qualitative cognitive interview data about the tools included in the prototype toolkit obtained in the previous Delphi study. In this study, which is part of a larger project, only the WHODAS 2.0 and P-scale were assessed on their aspects of validity, usefulness and relevance among the study population. For this purpose, we conducted a pilot and an initial cross-sectional validation of the tools among persons living with NTDs in Brazil, a country with a very vulnerable population for NTDs (Hotez, 2014).

3.3.2. Sampling method and study population

Participants for the pilot and initial validation were recruited through purposeful and convenience sampling**.** The sampling criterion was; ‘participants should have Chagas disease, (neuro) cysticercosis/taeniasis, leprosy, leishmaniasis or schistosomiasis and live in Ceará state of Brazil’. This state, situated in the upper-middle income country Brazil, is known for its very high burden of the included NTDs and their disabling and impoverishing consequences (World Bank, 2015). Participants were easily accessible through the close contact that local principal investigators have with health care workers from two hospitals; the Hospital Universitário

Walter Cantídío UFC and the Hospital São José de Doenças Infecciosas. Here, participants were identified by checking medical records from outpatients visiting departments such as neurology (persons with Leishmaniasis may have epilepsy), cardiology (persons with Chagas disease may have heart problems) and dermatology (e.g. leprosy and leishmaniasis). Consequently, many participants could have morbidity and disability related to a chronic form of the concerned NTD.

Inclusion criteria

The medical records were checked on the following inclusion criteria:

- Being diagnosed with Chagas disease, (neuro)cysticercosis/ taeniasis, leprosy, leishmaniasis or schistosomiasis;
- Having at least one of the clinical NTD-related characteristics as listed in Appendix 7
- Being 18 years or above;
- Living in Brazil;
- Being able to answer questions during an interview;
- Willing to give informed consent (written or verbal).

Patients diagnosed very recently were excluded from participation since it was expected that these patients have less or no disabling NTD conditions. To increase the diversity among the included participants, we aimed to recruit males and females, persons aged between eighteen and ≥ 85 years and persons living in differing geographical environments (rural versus urban etc.).

Sample size

Despite the fact that this study used purposive sampling; it was not directly preferable to base our sample size on theoretical saturation as is usual in such samples (Guest, Bunce & Johnson, 2006). Instead, we aimed to recruit around ten persons per NTD. This method fitted the purpose of the current study-objective better and was expected to represent more variation within the sample. Consequently, focusing on five of the NTDs endemic in Brazil, it was estimated to recruit a total sample size of 50 persons.

3.3.3. Data collection

Quantitative and qualitative cognitive interview data about two prioritized tools, the WHODAS 2.0 and P-scale, were collected using interview-administered versions of both, and by asking one question after each tool and three cognitive questions after the selected tool from the suggested toolkit. Prior to the data collection, a pilot study was conducted among thirteen participants. This pilot aimed to practice with the tools, to evaluate the comprehension, indicate time needed for administering, and the sequence in which the tools were best used. Following this, the initial validation took place, from June 22 until July 9, 2015, administering the tools and cognitive questions to investigate their ability to detect morbidity and disability across the included NTDs. This was done by four trained researchers. Besides this, participants were also asked to provide economic demographic information, mostly based on the demographic and background information questions from the WHODAS 2.0 (Üstün, 2010).

WHODAS 2.0 (12-items)

Functional limitations due to the included NTDs were assessed by the generic WHODAS 2.0 (Üstün, 2010) (Appendix 8). For this study the 12-item version of the WHODAS was used. This shortened version was chosen, because it saves time, while it is still reliable in assessing the same domains as the 36-item version (Andrews et al., 2009). This is expected to benefit the usability of the toolkit when users want to administer more than one tool and still want to obtain a good indication of the respondents' function status. The six life-domains the WHODAS 2.0 assesses are 1: cognition, 2: mobility, 3: self-care, 4: getting along, 5: life activities and 6: participation (Üstün, 2010). Response options range from 1: no difficulties to 5: extreme or cannot do, referring to the past 30 days. Since the WHODAS 2.0 was already validated widely (Cheung et al.,2014; Tazaki et al., 2014; Federici et al., 2009; Silva et al., 2013) in at least 27 languages (Üstün, 2010), the assessment of aspects of validity, usefulness and relevance for the target groups was the most important. The following question was asked after administering the WHODAS 2.0: *"This second set of questions was developed to get a better understanding of the difficulties people may have due to their health condition. Were these questions relevant/ meaningful for you? Why?"* Statistical analysis of the outcomes of the WHODAS 2.0 was not useful in this study, given the number of participants per NTD. Only the 'simple scoring' was used (Üstün, 2010).

P-scale (Short)

Participation restrictions due to the included NTDs were assessed by a shortened Portuguese version of the P-scale (Appendix 9). The inclusion of this shortened version was preferable because it saved time, asking thirteen questions instead of eighteen. The P-scale is initially developed to assess needs for rehabilitation of persons with leprosy, disabilities and other diseases that may restrict participation and/or cause stigma (version 4.2) (Van Brakel et al., 2006). The P-scale contains eight ICF-domains related to participation: learning and applying knowledge, communication, mobility, self-care, domestic life, interpersonal interactions, major life areas, and community, social and civic. Participants are asked if they perceive the level of a particular aspect of participation is the same as that of a peer. If not, they are asked whether they perceive that as a problem and if so how big this problem is; 1: no problem, 2: small, 3:

medium or 5: large. Results from the systematic literature review showed that there is already positive information on the cultural-equivalence of the P-scale (Kelders et al., 2012; Van Brakel et al., 2006; Jansen 2012; De Zeeuw et al., 2014). However, the review showed less information about what the target group thinks of the P-scale. In this study, this question may be answered by asking the following question after administering the P-scale; *"This third set of questions was developed to identify possible participation restrictions in the society. Were these questions relevant/ meaningful for you? Why?".*

Cognitive-data

To obtain information about aspects of validity and what the NTD target groups think about the relevance of the toolkit-tools, the following three cognitive questions were asked after administering all the tools;

1. *"Considering all the sets of questions we went through, do you consider that the total duration for the interview was acceptable / important, considering the (expected?) improvement of care?"*
2. *"Considering all the sets of questions we went through, in general terms, do you consider the usefulness positive or negative? Do you have any specific comments?"*
3. *"Is there something else that you want to tell us about the issues we talked about but were not included in this interview? Do you have any specific comments?"*

Besides this, the interviewers estimated the participants' comprehension of the tools by indicating and reporting when there was a need for examples and, additionally, when rephrasing of a question was necessary.

3.3.4. Data management and analyses

After conducting the interviews, quantitative data from the tools were entered into customized forms using Epi-Info version 7. Data obtained with the WHODAS 2.0 and the P-scale Short, were analyzed with simple statistical methods; calculation of total-scores per NTD, calculation of question related total scores, comparison of experienced morbidity and disability scores between NTDs and determination of overall comprehension of the questions. Qualitative data from the answers on the cognitive questions were recorded directly and participant names were converted to numbers. Concurrent with the collection of data, we transcribed and translated the Portuguese recordings into English. At the end these transcripts were analyzed using qualitative analyzing methods such as open coding, axial coding and selective coding. This analysis aimed to obtain a first impression concerning aspects of validity, usefulness and relevance of the WHODAS and the P-scale within the toolkit.

The validation study was submitted to the Ethics Committee (CEP) of the University Hospital Walter Cantídio, Federal University of Ceará (UFC). At the start of each interview, the participant was informed about the study and his/her rights (privacy, restricted use of information, etc.) related to it and asked for informed consent. Information the participants provided and information obtained from medical records was handled with integrity and confidentiality and was only accessible for directly involved persons.

4. RESULTS

4.1. Phase 1: Results from the Delphi study

(Sub-questions 1 and 2)

The first phase of this study answered sub-questions 1 and 2 with the purpose of collecting priority areas and tools aimed to asses and monitor NTD-related morbidity and disability, presenting a prototype toolkit, and deciding tools to be validated in the Ceará state of Brazil. Both sub-questions are answered by means of the Delphi study as explained in the methods.

4.1.1. Numbers of participants and countries represented

Table 3. Numbers and percentages on participation within the Delphi study
The first column lists the rounds this Delphi study comprised; the second column lists the total of invited experts per round and the third column shows the total number and percentage of completed e-surveys per round.

Action	Total invited	Total completed e-surveys (%)
Invited for & introduced to Delphi study	30	n/a
E-survey 1 (first round)	40	19 (47.5)
E-survey 2 (second round)	24	15 (62.5)
E-survey 3 (third round)	22	12 (54.5)

Table 3 shows that an initial number of 30 key-informants were invited to participate in the Delphi study. Twenty of the 30 invitees confirmed participation and therefore received the link to the first e-survey. Reasons to reject participation were “being based in an isolated area” and “not being an expert”. Additionally, another twenty key-informants, who were introduced by persons that were invited previously, also received the first e-survey link. Twenty-one out of the total of 40 persons that received the first e-survey started the e-survey, two of these 21 were incomplete. A link to the second e-survey was sent to 24 key-informants of whom 21 were experts participating in the first e-survey and three were new participants. This second e-survey was started by eighteen persons. However, three questionnaires were not completed. The third e-survey was sent to 22 experts of whom twelve completed the questions. Countries presented in this study included Australia, Brazil, Ethiopia, Ghana, Indonesia, India, Myanmar, Nepal, Netherlands, Nigeria, Norway Paraguay, Togo and the United Kingdom. Many of those experts have been working in NTD endemic countries.

4.1.2. Results from the first e-survey: investigating priority areas and tools

(First sub-objective of the study)

Field of expertise

Figure 6 shows the distribution of NTDs Delphi participants were involved in. All 21 participants indicated their field(s) of expertise and together they selected nine differing NTDs as being their field of expertise. The experts identified the option ‘other’ to be mental health and stigma. The results show that half (9) of the eighteen NTDs were represented in this study. Most NTD work-experience proved to be leprosy (13 experts), LF (7 experts) and trachoma (6 experts).

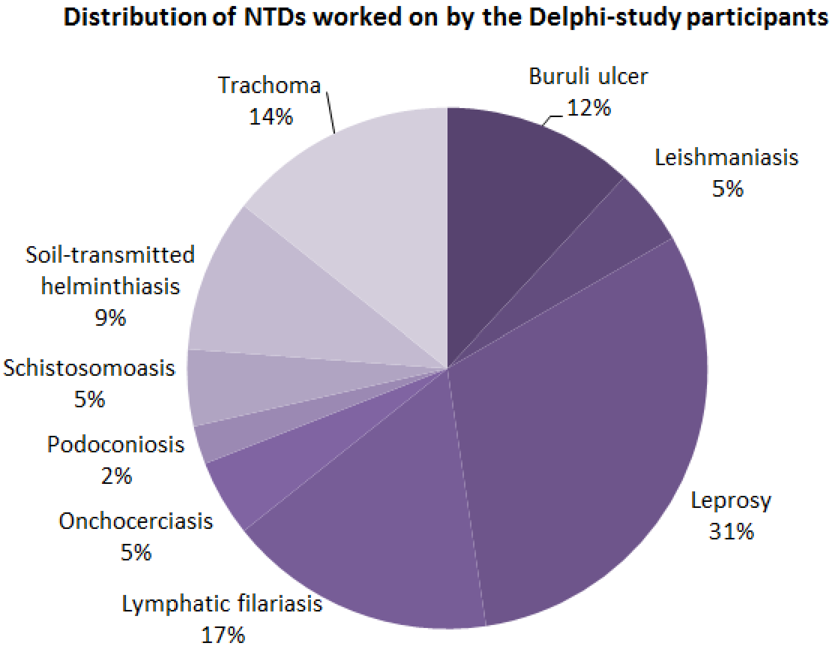


Figure 6. Distribution of the nine NTDs Delphi participants were involved in.
The 21 participants had the possibility to indicate more than one area of expertise, resulting in 42 entries.

Most important ICF domains

All 21 experts (who started the survey) selected one or more ICF domain(s) as being important in the assessment of NTD-related morbidity and disability. Eighteen experts stressed the importance of the concepts ‘health condition’, ‘impairments’, ‘participation restrictions’ and/or ‘environmental factors’. Sixteen experts (also) stressed the importance of NTD-related ‘functional limitations’ and fourteen experts (also) emphasized the importance of ‘personal factors’.

Collection of currently used and/ or potential tools

Tools that were currently used and/or being indicated as having potential for assessment, are listed in Table 4.

Table 4. Currently used and/or potentially important tools for assessment and monitoring of NTD-related morbidity and disability
The first column lists the ICF-domains and the second column lists the currently used tools among 19 Delphi study participants divided over the differing domains. The third column summarizes how often the listed tools are mentioned.

Domain of assessment and monitoring (ICF domain)	Tool	Times mentioned
Impairments **(Body functions and body structures)**	WHO Disability Grading	6
	Eyes Hands Feet (EHF) score	5
	Mental health (depression): Patient health Questionnaire (PHQ) 9	1
Functional limitations **(Activity)**	Screening of Activity Limitation and Safety Awareness (SALSA)	8
	WHO Disability Assessment Schedule (WHODAS) 2.0	1
	Buruli Ulcer Functional Limitation Scale (BUFLS)	3
	WHO Quality of life (QoL): WHO-QoL BREF	1
Participation restrictions **(Participation)**	Participation Scale (P-scale)	9
	WHO Disability Assessment Schedule (WHODAS)	1
	Stigma Assessment and Reduction of Impact (SARI) scale	1
	Jacoby Scale	1
	WHO Quality of life (QoL): WHO-QoL BREF	1
Environmental factors	Explanatory Model Interview Catalogue (EMIC) individual	1
	Explanatory Model Interview Catalogue (EMIC) community	1
Personal factors	-	-

This fourth question was answered by twenty experts, of which thirteen provided information about tools that could be found on the internet. The most frequently cited tools were the Screening of Activity Limitation and Safety Awareness (SALSA) scale, the Participation-scale (P-scale), the WHO Disability Grading and the Eyes Hands Feet (EHF) score (a leprosy specific tool). These tools were not only mentioned by leprosy experts who represented 31 % of the NTD experts, the SALSA and the P-scale were also suggested by LF and Buruli ulcer experts and a trachoma expert suggested the WHO disability grading. The WHODAS was mentioned under both of the domains it assesses; activity and participation. The same applies to the WHO Quality of life (WHO-QoL BREF). Results show that, beside the Explanatory Model Interview Catalogue (EMIC), there were not any tools suggested for assessment and monitoring of environmental factors. However, experts mentioned some factors that might be important to assess in this domain: environmental hygiene, community services, accessibility of services, quality of services, accessibility to support, economic support provisions and barriers. No suggestions were given for the assessment and monitoring of personal factors. Answers on the sixth question in this e-survey: *"Do you have any knowledge of tools that are used for other NTD(s) or disabilities* *that you believe could be important in the assessment of different aspects of NTD-related morbidity and disability in your field?"* did not add any additional new tools.

Interventions to facilitate solutions for Morbidity Management and Disability (MMD)

Ideas for morbidity and disability reducing interventions for NTD-affected persons were provided by twenty of the twenty-one experts, answering the seventh question of the survey: "*What do you think are the best interventions to facilitate solutions for morbidity management and disability prevention?"* It may be possible to use these ideas for the development of a guide or manual for the generic cross-NTD toolkit. Figure 7 provides an overview of possible interventions (as can be seen in the inner circle) and their requirements (as can be seen in the outer circles).

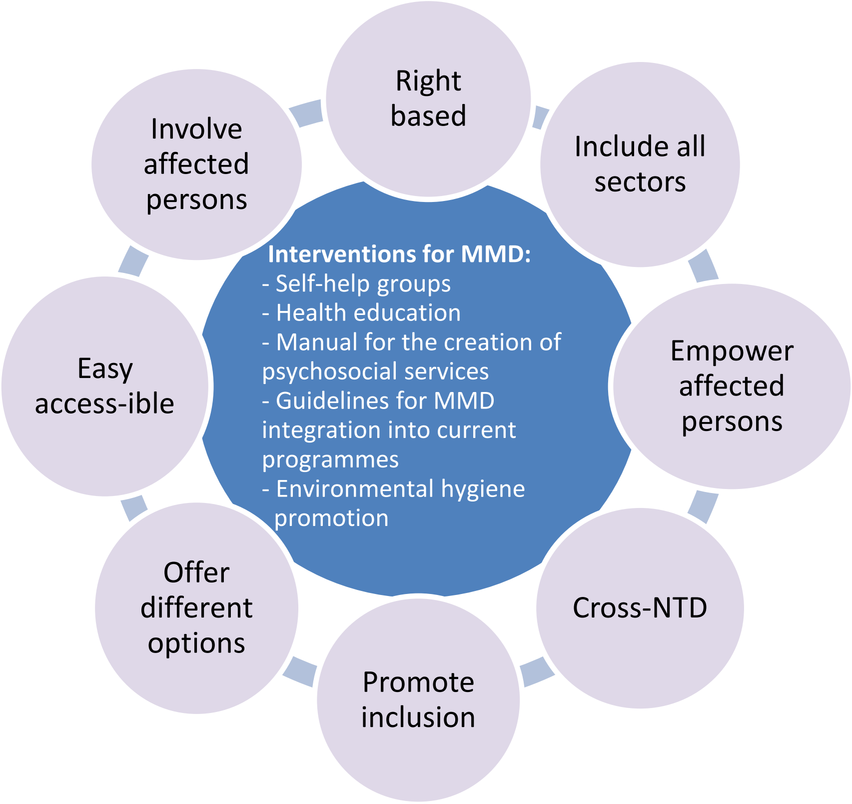


Figure 7. MMD-facilitating interventions mentioned by experts from the Delphi panel
The inner circle lists interventions for MMD and the outer circles shows the requirements for these interventions as collected from the experts answers.

4.1.3. Results from the second e-survey: prioritizing the tools

(First sub-objective of the study)

Tools for assessment and monitoring of NTD-related impairments

Concerning the ICF-domain of body functions and structures, one tool, the WHO disability grading & the EHF-score, was mentioned several times in the first e-survey. Therefore, this tool was included together with the WHO ICF-checklist and the Buruli ulcer patient's prevention of disability (POD) assessment that were obtained through the literature search (Zeldenryk et al., 2012) and directly from one of the principle researchers (Van Brakel). Table 5 reveals that the WHO ICF-checklist was by far the most frequently mentioned tool for assessment of NTD-related impairments (13 times). The other two, more specific tools, were mentioned less often. It can also be seen that nobody ranked a tool in the third place and only four participants ranked a tool in second place.

Table 5. Ranking results of tools for the assessment and monitoring of NTD-related impairments
All fifteen experts ranked (second and third column) tools (first column) for the assessment and monitoring of NTD-related impairments.

Tool	Ranked 1st	Ranked 2nd	Ranked 3rd	Total score
WHO disability grading & the EHF-score	4	-	-	**12**
WHO ICF-checklist	9	4	-	**35**
Buruli ulcer patient's prevention of disability (POD) assessment form	2	-	-	**6**

The total scores as displayed in the fourth column are obtained by awarding points based on ranking place (1st = 3 points, 2nd = 2, 3rd = 1 (not mentioned here) and not ranked = 0).

Tools for assessment and monitoring of NTD-related functional limitations

Within the ICF-domain of activity, a total of five tools were listed. The SALSA scale and the WHODAS 2.0 were both adopted from results from the first e-survey. The other three tools, the Green Pastures Activity Scale (GPAS), the Barthel Index (BI) and the Washington Group Questionnaire on disability were suggested by the experts (Van Brakel et al., 1999; Tsutsumi et al., 2007; Riewpaiboon et al., 2009; Miller et al., 2011; Bachani et al., 2014). Table 6 shows that the total scores for these tools were very close, except the total score for the BI with only eight points. The SALSA was ranked first by five experts while other tools were ranked first by only three or fewer experts. The WHODAS 2.0 however received the highest score (22 points) followed by the GPAS (20 points) and the Washington Group Questionnaire on disability (nineteen points).

Table 6. Ranking results of tools for the assessment and monitoring of NTD-related functional limitations
Fourteen experts ranked (second, third and fourth column) tools (first column) for the assessment and monitoring of NTD-related functional limitations.

Tool	Ranked 1st	Ranked 2nd	Ranked 3rd	Total score
Screening of Activity Limitation and Safety Awareness (SALSA) scale	5	-	2	**17**
WHO Disability Assessment Schedule (WHODAS) 2.0	3	6	1	**22**
Green Pastures Activity Scale (GPAS)	3	5	1	**20**
Barthel Index (BI)	1	1	3	**8**
Washington Group questionnaires on disability	3	3	4	**19**

The total scores as displayed in the fifth column are obtained by awarding points based on ranking place (1st = 3 points, 2nd = 2, 3rd = 1, and not ranked = 0).

Tools for assessment and monitoring of NTD-related participation restrictions

Results from the first e-survey showed six tools that could be used within the ICF-domain of participation. Although most of these tools indeed linked to this domain, two exclusively focused on assessing stigma. Tools that were included after exclusion of the stigma-focused tools are listed in Table 7, showing their ranking scores. The P-scale was obtained from the first e-survey and the Impact on Participation and Autonomy (IPA), the London Handicap Scale (LHS) and the Sheehan Disability Scale (ShDS), were all found in literature (Stevelink & Van Brakel, 2013. Sudhir et al., 2012.). The P-scale was ranked first by eight experts and received a total score of 30 points followed by the IPA (21 points), LHS (12 points) and the SDS (3 points).

Table 7. Rankings of tools for the assessment and monitoring of NTD-related participation restrictions
Twelve experts ranked (second, third and fourth column) tools (first column) for the assessment and monitoring of NTD-related participation restrictions.

Tool	Ranked 1st	Ranked 2nd	Ranked 3rd	Total score
Participation scale (P-scale)	8	3	-	**30**
Impact on Participation and Autonomy (IPA)	4	4	1	**21**
London Handicap Scale (LHS)	1	4	1	**12**
Sheehan Disability Scale (ShDS)	-	1	1	**3**

The total scores as displayed in the fifth column are obtained by awarding points based on ranking place (1st=3 points, 2nd= 2 points, 3rd=1 point and not ranked=zero points).

Tools for assessment and monitoring of NTD-related environmental factors

Respondents to the first questionnaire did not suggest any tools for direct assessment and monitoring of environmental factors. The EMIC tools (individual and community version) which were mentioned are indeed part of environmental factors, however, more specifically focused on stigma. Therefore, in the second e-survey, we suggested the Craig Hospital Inventory of Environmental Factors (CHIEF) scale as a possible tool to assess and monitor environmental factors (Harrison-Felix, 2001; Van Brakel & Officer, 2008). Participants were asked to give their opinion regarding the suitability of the CHIEF and to suggest any other instruments they knew that might be used to assess and monitor environmental factors across NTDs. This question was answered by seven experts. Five thought it was interesting and relevant to include the CHIEF. Two experts said they would not use the CHIEF because it adds little information to the toolkit and was not developed for low-income countries. One other tool suggested was the Quality of Care as seen through the Eyes of the Patient (QUOTE) for persons with tuberculosis (Van Campen et al., 1998).

Tools for assessment and monitoring of NTD-related personal factors

Results from the first e-survey also did not directly suggest tools that were linked to personal factors, yet some aspects of personal factors were actually mentioned, and the experts also named tools that could be used to assess these aspects such as the WHO-QoL BREF for the assessment of quality of life (QoL). The Rosenberg self-esteem scale was found in literature and included in this question because of the relevance of measuring self-esteem. Experts were asked to select the tool that, in their opinion, should be included in the toolkit (Table 8). The WHO-QoL BREF received the highest total score (21 points).

Table 8. Rankings of tools for the assessment and monitoring of NTD-related personal factors

A third (ten) of the experts ranked (second column) tools (first column) for the assessment and monitoring of NTD-related personal factors.

Tool [aspect]	Ranked for inclusion	Total score
WHO Quality of Life (WHO-QoL) BREF [quality of life]	7	**21**
Rosenberg self-Esteem scale [self-esteem]	5	**15**
“I do not think a specific tool to assess these areas is necessary”	3	**9**

The total scores as displayed in the third column are obtained by awarding points based on inclusion; yes = 3 points or no = 0 points.

NTD-specific tools

At the conclusion of the e-survey, participants were asked to indicate the need for inclusion of NTD-specific tools in the cross-NTD toolkit. Leprosy-specific tools, such as the EHF score, were mentioned four times. A similar tool was suggested for LF, and a BU-specific tool was also provided by one of the experts. One of the experts advised against the inclusion of NTD-specific tools, because this may cause a “*false sense of assessment and completeness*” (Delphi study participant, 2015). Therefore, at this stage NTD-specific tools will not be included in in the toolkit.

Prototype toolkit

Following the result from this second e-survey, it was possible to propose a prototype toolkit (Table 9).

Table 9. Prototype toolkit, based on highest total scores per domain
The first domain lists the five ICF-domains and the additional domains on stigma and mental health. The second column summarizes the tools that received the highest score (third column) within their domain or were included for other reasons as described in the fourth column.

Domain	Proposed Tool	Total points received	Justification inclusion
Body functions and structures	WHO ICF Checklist	35	Highest priority
Activities	WHODAS	22	Highest priority
Participation	P-scale	30	Highest priority
Environmental factors	CHIEF	-	Five experts thought it was interesting and relevant to include the CHIEF
Personal factors/ Quality of life	WHO-QoL BREF or Dis.	21 (for BREF version)	Disability version was included later and proposed as a second option.

4.1.4. Results from the third e-survey: tools to be preliminary validated

(Second sub-objective of the study)

Tools prioritized for the initial validation

From the prototype toolkit as given in Table 9, twelve of the twenty-two invited experts prioritized tools that they thought should be validated. Table 10 shows the top four tools, 1: WHO ICF Checklist (34 points), 2: WHODAS (26 points), 3: WHO-QoL BREF or Dis (17 points), 4: P-scale (14 points). Participants were also asked to indicate which of the QoL-tools they considered most suitable to include in the toolkit; the BREF version or the more disability-focused Dis version. A large majority (nine experts) chose the WHO-QoL Dis.

Table 10. Tools to be initially validated based on highest total scores
All twelve Delphi study participants ranked the tools, as proposed in the prototype toolkit, for the initial validation.

Tool (Field of assessment)	Ranked 1st	Ranked 2nd	Ranked 3rd	Ranked 4th	Total score
WHO ICF Checklist (Impairments)	8	-	1	-	**34**
WHODAS (Functional limitations)	2	4	3	-	**26**
P-scale (Participation restrictions)	1	2	-	4	**14**
CHIEF (Environmental factors)	-	-	2	1	**5**
WHO-QoL BREF or Dis (Quality of Life)	-	4	2	1	**17**

The total scores as displayed in the sixth column are obtained by awarding points based on ranking place (1st = 4 points, 2nd = 3, 3rd = 2, 4th = 1 and not ranked = 0).

4.2. Phase 3: Results from the initial validation of the WHODAS and P-scale Short (Sub-question 4)

The third and last phase of the study comprised the collection of data of approximately 50 persons living in the Ceará state of Brazil with NTD-related morbidity and disability. This phase answered the fourth sub-question of the study through an initial validation of the toolkit and asking cognitive questions after each tool and the toolkit as a whole to describe aspects of validity, usefulness and relevance of two of the tools (the WHODAS 2.0, 12-items and the P-scale Short).

4.2.1. Sample size and characteristics

The WHODAS 2.0 (12-items) and the P-scale Short (and other) tools were validated among persons with Chagas disease, leprosy, leishmaniasis (visceral) or schistosomiasis (Table 11). From five participants it was known that they had possible disabling comorbidities such as schizophrenia, alcoholism, HIV, Diabetes Mellitus or aphasia. Most (8) of the leishmaniasis patients were recruited in the Hospital São José de Doenças Infecciosas; the other patients were seen in the Hospital Universitário Walter Cantídío UFC. It was impossible to recruit enough participants with (neuro) cysticercosis and therefor the single person with this NTD is excluded from data analyses. Except for schistosomiasis, for which only four female participants were found, each included NTD was present in ten participants. Thirty-four persons, 18 male and 16 female participated.

Table 11. Socio-demographic characteristics participants

	Chagas disease n=10		Leprosy n=10		Leishmaniasis n=10		Schistosomiasis n=4	
	♂ n=6	♀ n=4	♂ n=5	♀ n=5	♂ n=7	♀ n=3	♂ n=0	♀ n=4
Age (mean)	62	50	51	43	46	47	-	62
Education (n)			*					
Illiterate	1					1		1
Literate	1	1			2			1
Attended primary school	2	2	3	2	2	1		1
Attended secondary school	2	1	1	3	2	1		
Attended higher education					1			1
Employment (n)			*					
Paid work	2			2	1			2
Self-employed	3	2	1		1			
Keeping house/ homemaker		1	1			2		1
Retired	1				3	1		1

	Chagas disease n=10		Leprosy n=10		Leishmaniasis n=10		Schistosomiasis n=4	
Unemployed (health reasons)		1	1	2	1			
Unemployed (other reasons)				1				
Other					1			

N = 34
* Data concerning education and employment is missing for one participant.

On average, participants with Chagas disease or schistosomiasis were ten years older than participants with leprosy or leishmaniasis. All leprosy patients attended primary school or secondary school, while the other groups also included persons who are illiterate or literate. One of the participants completed higher education, another one started but did not complete. Of the nineteen persons that were able to work and not retired, fourteen persons had paid work or were self-employed. Five persons said that they could not work because of health reasons.

4.2.2. Quantitative study results

Table 12 lists the NTD-related descriptive information about functional limitations (assessed with the WHODAS 2.0) and participation restrictions (assessed by the P-scale Short). Total scores for the WHODAS were obtained by using the 'simple scoring' method as described in the users-manual (Üstün, 2010). Total scores for the P-scale were obtained by a similar method; the scores assigned to each of the reply options were summed. The higher the scores, the more problems.

Table 12. Descriptive characteristics on the WHODAS 2.0 and P-scale Short
Mean total scores from the tools are given in the second and third column of the table and divided per (NTD) group and sex as is shown in the first column. Maximum obtainable score for WHODAS 2.0 was 60 points and for P-scale Short 65 points.

(NTD) group	WHODAS 2.0 (12-items) Mean total score (range)	P-scale Short Mean total score (range)
Chagas disease	**19**	**11**
Male (n=6)	17 (10 – 27)	13 (0 – 41)
Female (n=4)	22 (15 – 29)	9 (4 – 17)
Leprosy	**23**	**10**
Male (n=5)	23 (21 – 28)	10 (6 – 15)
Female (n=5)	17 (14 –22)	7 (0 – 15)
Leishmaniasis	**16**	**3**
Male (n=7)	18 (9 – 35)	4 (0 – 8)
Female (n=3)	11 (10 – 11)	1 (0 – 3)
Schistosomiasis	**20**	**15**
Male	-	-
Female (n=4)*	20 (12 – 36)	15 (1 – 35) *

(NTD) group	**WHODAS 2.0 (12-items)** Mean total score (range)	**P-scale Short** Mean total score (range)
Overall	**19**	**9**
Male (n=18)	20 (9 – 35)	9 (0 – 41)
Female (n=16)	18 (10 – 29)	7 (0 – 35)
Age group		
< 65 years (n=26)	25 (9 – 36)	11 (0 – 41)
> 65 years (n=8)	19 (11 – 35)	3 (0 – 21)

N = 34
*n=3 for P-scale

Test-results WHODAS 2.0 (12-items)

Participants with leprosy (n=10), especially the males, had the highest mean total score (23) on the WHODAS. Question-related total scores of all leprosy patients were the highest for question 5 ‘*emotionally affected*’ and Question 7 ‘*walking long distances such as a kilometre*’ (both 30). Following the leprosy group, a high level of problems in functioning were also reported in participants with schistosomiasis (mean total score was 20). The lowest NTD-related mean score applied to the four female participants with leishmaniasis (11 points). Interestingly, the highest individual total score (35 points) was also reported by a participant with leishmaniasis. Overall, question-related highest scores were almost similar to the scores for leprosy; fourteen participants reported to have at least moderate difficulties with emotional affection, thirteen with walking long distances and fifteen with standing long periods (Question 1). No limitations were experienced in maintaining a friendship (Question 11) and less limitations in dealing with unknown people (Question 10). Overall, males (n=18) reported slightly more problems in functioning than females (n=16). Mean scores were also higher for participants younger than 65 (n=26), differing 6 points on the total score with participants who were older (n=8). Highest and lowest scores (range) within those age-groups were approximately the same.

Test-results P-scale Short

Participation restrictions were highest for the three participants with schistosomiasis (mean total score was 15) and the ten participants with Chagas disease (mean total score was 11). However, the latter score which is only 1 point higher than the mean total score for participants with leprosy (10 points), is strongly influenced by the high individual score of 41 points that was reported for one of the participating males. Overall, highest scores and most difficulties were found for the first questions “*Do you have equal opportunity as your peers to find work?*”, and the second question *“Do you work as hard as your peers do?*” Three of the six males with Chagas disease answered to have “large” problems to this second question. Overall, lowest total scores on the individual questions applied to Question 7: "*Do you have the same respect in the community as your peers?*" and Question 13: "*Are you comfortable meeting new people?*", both received only 1 point. This result is in line with the comparable question about 'dealing with people you do not know' from the WHODAS. Also comparable with the WHODAS scores was that both male participants and participants who are younger than 65 years reported higher scores on the P-scale.

Duration and comprehension of the tools

For both the WHODAS and the P-scale Short, the duration of administration was recorded. The average time for administering the WHODAS 2.0 was 9 minutes (range 2 – 24). For the P-scale Short (13 questions) the average duration was 7 minutes (range 3 – 26). The participant who needed 24 minutes for the WHODAS interview had a hard time understanding the questions; twelve of the fifteen questions needed to be rephrased. Overall, comprehension was poor for WHODAS Question 5: *"How much have you been emotionally affected by your health problems?"* (examples were needed in four cases and rephrasing in eight cases). Question H3: *"In the past 30 days, not counting the days that you were totally unable, for how many days did you cut back or reduce your usual activities or work because of any health condition?"* needed examples in six cases and rephrasing in two cases. The concept 'usual activities or work' from question H3 raised questions from the participants. Other items for which some difficulties with comprehension were reported were 3: "*Learning a new task, for example, learning how to get to a new place?*" (examples were needed in three cases and rephrasing in four cases) and 6: "*Concentrating on doing something for ten minutes?*" (examples were needed in four cases and rephrasing in three cases). The P-scale Short seemed to raise fewer questions and was understood more easily than the WHODAS. Question 12, "*In family discussions, does your opinion count during important decision making?*", however needed examples in three cases and rephrasing in another three cases. A possible explanation for the lower comprehension of this question is that the concept 'discussion' was often interpreted as 'having a fight'. The peer concepts seemed to cause little difficulty; only one of the participants could not think of a peer to be compared with and became very emotional since all her peers passed away. Therefore the P-scale was not administered for this older lady with schistosomiasis.

4.2.3. Qualitative study results

Qualitative data were obtained by asking the participants if and why they thought the questions were meaningful. Regarding the WHODAS and P-scale Short, all 34 participants answered the validation questions positively. Three recurring reasons emerged. First, participants became more aware of their health condition and reported that the questions may also result in awareness in others (such as peers and health care workers). Second, especially the Chagas disease and leprosy patients emphasized that the questions stimulated verbal expression about their disease, something which is not very common. Third, in particular leprosy and leishmaniasis patients underlined the appropriateness of the questions.

The WHODAS 2.0 (12-items) for persons with Chagas disease, leprosy, leishmaniasis or schistosomiasis

Many participants in all four NTD-groups mentioned that they, due to the WHODAS-questions, had learned something or became more aware of their health-situation. The following quotes illustrate this awareness:

> "*Because you are opening my mind, my disease, we know more things, what I am able to do and what not."* (F, 61, Chagas disease)

> *"I understood what I feel, like this weakness, this fatigue, it is not only me."* (M, 40, leishmaniasis)

Participants also indicated that awareness could be preventive for further development of NTD-related disability. The following quotes illustrate the possible effect of their personal awareness and awareness in others:

> "*It is important because if I had felt* [the problems] *then it maybe could help me."* (M, 44, Chagas disease)

> "*Hopefully this will serve as an example for people and that people who get this disease run for curing, right? Because it took me so much time and that is why I am having these difficulties, you know?"* (M, 52, leprosy)

The last quote is from a participant who lost parts of his fingers because of a very late diagnose of leprosy.

Many of the participants with Chagas disease or leprosy indicated importance of the contribution of the questions to make them talk about the disease they are dealing with. One of the Chagas disease patients said that no one has ever asked such 'things' before. While participants in both groups indicated the importance of sharing feelings and daily-life experiences, they also emphasized the difficulties some people have with talking to other people about their disease:

> "*There are many people who do not like it or want it. It is a fear for people. Sometimes people get a disease and are afraid to talk."* (M, 81, Chagas disease)

> "*...it is very good for persons to talk with someone like that, right? To explain. You stay with those things* [feelings, thoughts, fear] *only with yourself, right? Because I only talk to my wife, only she, you know? Because she knows everything about my life, I know all about hers, we met in the hospital, right? I feel good about this, because my neighbours know nothing about my life, you see?"* (M, 58, leprosy)

The effect of the NTDs on a person's functioning, and therefore the relevance of the WHODAS for the assessment of disability is mainly indicated by participants from the leprosy group and the leishmaniasis group as illustrated in the following quotes:

> " *...because we try to live the most common way as possible, but our body requires other ways of living. There are times that you do not have* [...] *it is* [...] *if you want to move the way you want, than you end up having to adapt many things, to change the way of doing a task or having a relationship, sometimes even to work..."* (F, 36, leprosy)

> *"So, you think it* [life?] *is like a book isn't it? So, it is open, you think you can do everything and at the time you try you are not able to do it, but you try. Like in my case, I try but I am never able to because of pain that disturbs me very much. I keep trying but it is never good, so I gave up. Brazilian people never give up, but for me, I cannot do, I tried until my limits but I could not. Thus, for me these questions are really important."* (F, 23, leprosy)

> *"It was important because, in day-to-day things I* [...] *use a lot of patience to do what is impossible, however necessary. I had a few hours I mistreated myself a lot, even cry, I cry right? It is bad when you feel like doing something and cannot do it anymore, isn't it? Because of the disease, you see?"* (M, 74, leishmaniasis)

The P-scale Short for persons with Chagas disease, leprosy, leishmaniasis or schistosomiasis

One of the main reasons for positive reactions on the P-scale Short was the appropriateness of its questions for especially persons with Chagas disease, leprosy and leishmaniasis. Especially persons with leprosy confirmed the effect of leprosy on participation and the possible participation restrictions:

> Interviewer: "*You commented that you are closed* [inside]*, has it to do with the disease or not, have you always been like that?"*
> Participant: "*No, it was after leprosy, because when I had my crisis, I was part of a church, when I arrived at the church no one sat next to me, because I was full of injury, not washed, smelled too, since I could not get perfume, I had to bathe only with sulphur soap. It made me insecure; from the beginning until now I am like this. Before I was more open, I will not lie* [...] *it changed my way of being."* (F, 36, leprosy)
>
> *"I am very afraid of them knowing I have leprosy, right? I have a lot of fear, prejudice still exists.* […] *I have a house at the shadow side and was very afraid that they [the neighbours] would discover and spread the news. The prejudice will not be over I think. I would not go to places if people knew."* (*M, 58,* leprosy)
>
> *"Yes, because* […] *there are people we know who have Kala-azar isn't it? We do not want to* [...] *think that you get it [Kala-azar] from another person heh? But it is not contagious isn't it? One time when I stay at the clinic a woman asked: "What treatment do you get?" The Kala-azar! "Wow, you had Kala-azar?" I had, I am standing here, so good nowadays, telling the story. Some persons do not want to sit closely that person who had that disease. I do not know. I think that is horrible*." (F, 34, leishmaniasis)

The influence of the participation restrictions also seemed to result in persons having difficulties and even fear of talking about their disease. Therefore, many participants, especially the ones affected by Chagas disease or leprosy, seemed grateful for the opportunity to talk (freely) about their day-to-day life and relationships, stimulated by the P-scale questions. The following quotes emphasize both the importance and difficulties of talking with someone about your life when having Chagas disease or leprosy:

> "*It was important. Because I am sharing my day-today life, my relationships, it is better for your health."* (M, 69, Chagas disease)

> "*It is very important, right, that* [...] *we talk with a person, right, about our life? Because, with the neighbours I cannot talk about it."* (M, 58, leprosy)

Participants in the three NTD-groups mentioned that the questions of the P-scale made them more aware of their health-situation. One of the Chagas disease patients even expressed his hope for universal awareness as the following quote illustrates:

> "*I liked it. Because I think that health is very poor* [in general], *poor treatment and an interview like this maybe the way for someone sees a way to improve our health."* (M, 59, Chagas disease)

5. DISCUSSION

The aim of this study was to develop a generic cross-NTD toolkit. The purpose of such a toolkit is to assess, monitor and evaluate NTD-related morbidity and disability aspects to: 1: identify priority areas for MMD services 2: monitor and evaluate the effect of interventions for preventing and managing disability and 3: collect data for advocacy purposes. The most important phases in the development were a Delphi study among NTD-experts and the initial validation of the toolkit among persons living with Chagas disease, leprosy, leishmaniasis or schistosomiasis in the Ceará state of Brazil. As part of a larger study, this study focused on the WHODAS for the assessment of functional limitations (Üstün, 2010), and the P-scale Short for the assessment of participation restrictions (Van Brakel et al., 2006).

5.1. Discussing inclusion of the WHODAS 2.0, 12-items

Results from the initial validation seem to indicate that the WHODAS could be relevant among the NTDs included in the study. Positive experiences among the study population were mainly found in the reported relevance of the WHODAS, the awareness the WHODAS raised and the open communication as stimulated by the questions.

Regarding the quantitative validation results, persons with leprosy received the highest mean total scores and persons with leishmaniasis the lowest. Previous studies also demonstrated high levels of limitations in functioning among persons with leprosy (Brouwers et al. 2012; Van Brakel et al., 2012). The study from the SALSA Collaborative Study Group (2007) sought ways to assess activity limitations among persons with leprosy and diabetes, and developed and validated the Screening of Activity Limitation and Safety Awareness (SALSA) scale in five differing countries including Brazil. In their study, adjusted for age and impairment level, SALSA scale scores were higher in Brazil (Salsa Collaborative Study Group, 2007). Also Van Brakel et al. (2012) who used the SALSA scale, mentioned that more than half of the leprosy affected study population in Indonesia reported to have activity limitations. A study among leprosy affected persons in The Netherlands used the self-administered version of the WHODAS 2.0 (36-items) among 82 responders and mentioned substantial high levels of activity limitations as well (Slim et al., 2010). No relevant publications were found focusing on or explaining the lower level of activity limitations due to leishmaniasis. Their low score may be explained by the less severe signs of visceral leishmaniasis in this group and the fact that this group had the lowest mean age.

Contrary to expectations, it was observed that the lowest total scores in the leprosy group related to Question 10 '*dealing with people you do not know*' and 11 '*maintaining a friendship'.* This finding seemed unexpected since persons with leprosy often experience participation restrictions and barriers in relationships (Van Brakel et al., 2006; Peters et al., 2014). Similar proportions between question-related outcomes were found in the other NTD groups. The lower score on these two questions might be explained by their less practical character compared to the very practical questions about walking, standing and household responsibilities in which problems are more easily detected. Overall results indicated that it is suitable to assess functional limitations within the current study population, especially among leprosy patients. This can be done with the WHODAS or the more activity focused SALSA scale as is often used in previous research.

Furthermore, inclusion of the WHODAS is also supported by the fact that the questions were reported not to be too challenging to this study population. A little surprising since these questions are relatively abstract in our opinion. This finding is in agreement with other Brazilian studies using the WHODAS 2.0 among pregnant women (Selveire et al., 2013) and patients with musculoskeletal pain (Silva et al., 2013). Analysing and comparing answers on the question about school/ work however was more complex, since there were fifteen participants who did not work and results were missing for three others as well. A similar situation was found in the study from Silva et al. (2013) among patients with musculoskeletal pain. Therefore it is advised to assess the results on this question separately. It is hypothesized that these scores would have been even higher if the participants who were unable to work had the option of indicating this inability.

5.2. Discussing inclusion of the P-scale Short

The most important reasons for the inclusion were the appropriateness of the measurement of participation restrictions experienced by people with NTDs, especially persons with leprosy. Such a result was expected since the original P-scale was also developed and found relevant and valid among persons with leprosy in Brazil (Van Brakel et al., 2006).

The usefulness of the P-scale Short, as part of the generic toolkit, was also found in the direct answers resulting from administering the tool. Similar to the functional limitations, participants with leishmaniasis scored lowest on participation restrictions. Here, on average this leishmaniasis group scored at least three times lower than participants with other NTDs. An explanation for this difference might be found in the less severe visceral leishmaniasis signs, the sampling variation due to the small sample size and non-random selection. Highest levels of participation restrictions were obtained for the small schistosomiasis group in which only three females completed the P-scale Short. Beside the possible explanation of the NTD-characteristics and signs and the very small sample, it remains difficult to find a clear literature-based explanation for this outcome. One inter-disciplinary study among mothers engaged in domestic work in Sudan even describes that there was no effect of the disease on social activities at all, making the high score in the current study even more questionable (Parker, 1992). A final explanation might be the relatively high age of the participants within this group. The second highest score among Chagas disease is probably mainly influenced by one individual, with a very high score. It may very likely be that, without this score, leprosy would have been the group with second highest experience of participation restrictions. This assumption could be supported by the fact that the original P-scale was developed for people with leprosy, who often experience high level participation restrictions (Kopparty, 1995; Withington et al., 2003) and by results from a P-scale validation among persons with leprosy in India, Nepal and Brazil (Van Brakel et al., 2006).

The usefulness of the P-scale Short can be supported even more by the decision to include this short version which means a reduction of at least 13 minutes on average of the administration time (the overall average administration time of the original scale was 20 minutes, the average time of the short version within this study was around 7 minutes)(Van Brakel et al., 2006). This time-reduction is surprising since only five of the eighteen questions of the original scale were omitted from the original scale. Also the high comprehension of the P-scale among the participants supports the inclusion. The 'peer'-concept, subject of discussion during a validation of the scale in Indonesia (Kelders et al., 2012), did not seem to cause many problems in the current Brazilian study. Only one of the participants was not able to answer because of the 'peer'-concept. This was not because of misunderstanding of the concept, but because she did not have a peer. Previous validation of the Brazilian-Portuguese version of the P-scale also reported minor problems with this peer-concept and recommended some strategies to overcome possible misunderstandings (Van Brakel et al., 2006). However, the use of the word 'discussions' in Question 12, did seem to cause some misinterpretations among the Brazilian population. While such a misinterpretation was not mentioned in previous validation of the P-scale in Brazil (Van Brakel et al., 2006), this questions was omitted during the validation of the scale in Indonesia in order to simplify the scale for this population (Kelders et al., 2012). In the current study, the misinterpretation of the word 'discussions' possibly caused the remarkably low question-specific total score on this question. In fact, some of the participants mentioned they interpreted the word ‘discussion’ as ‘having a fight’. As a result it could be assumed that there were more persons in this study population who may have misinterpreted this question and therefore answered that they did not have such problems. Brazilian researchers in the study recommend adapting this question to: *“In family decisions, is your opinion important?”*.

5.3. Discussing the study process

During the collection of the tools by answers from the Delphi panel, it was noticed that not all tools were correctly classified in the right ICF-domains. Besides, widely used tools such as the SDS (Parrillo & Donoghue, 2005) were ranked relatively low compared with less well known tools, here the EMIC (Weiss et al., 1992). Therefore it is suspected that the NTD-experts in this study were more familiar with the EMIC, which was used among NTDs represented by almost half of the Delphi panel (Stientstra, 2014; Stevelink et al., 2011; Rensen et al., 2011) than with the SDS which is mainly used in the mental health field and only validated among persons with leprosy and not among persons with other NTDs (Peters, 2014). It might also be that our piloted introduction on the ICF and attachment of the tools, as was done in the e-surveys, was still not sufficient enough.

Convenience sampling, as used for recruiting participants for the initial validation, is often associated with bias since the selected participants may not be representative of the study population (Lunsford & Lunsford, 1995). At the same time this sampling method, which was used in two local hospitals, contributed to a higher prevalence of disability and morbidity in the current study population. Results from such a study population are expected to demonstrate that the tools could detect morbidity and disability in persons affected by an NTD, which was the purpose of the study.

5.4. Strengths and limitations

A limitation of the whole toolkit itself is that it is not always able to identify the cause of the disabilities mentioned. It cannot be determined with certainty whether the reported disabilities are a consequence of the NTD or comorbidity such as for example Diabetes Mellitus. Inclusion of health professionals who could obtain such information could be a solution for this. However, the toolkit is also aimed to be applicable by non-health professionals.

A limitation of the Delphi study was the absence of statistical methods, for example Kendall's coefficient of concordance (Kendall's' W) to assess agreement among the experts. While Ju & Jin (2013) emphasize the strength and importance of such nonparametric statistics in Delphi studies, the questions and answering instructions of the current Delphi study were not designed to use this in the most optimal way. Nonetheless, the inclusion of experiences and suggestions from key-informants by means of a Delphi study is a strength of this study; their information was extremely important in developing the toolkit and could be used in further development of the toolkit as well. Generalization of the results from the pilot and initial validation should be done with even more care. The present study is an initial validation in which the number of subjects recruited was not large enough to draw definitive conclusions concerning the relevance and suitability of the toolkit. Findings and experiences from this initial validation are an important start of validations of the toolkit among differing NTDs, in differing environments and cultures.

5.5. Recommendations

In future research, it may be relevant to test the SALSA scale for possible inclusion in the toolkit, instead of the WHODAS. The SALSA scale was proposed several times in the first e-survey, is advised for use in combination with the P-scale by Van Brakel et al (2006) and American Leprosy Missions (2015) proposed the scale for the identification and measuring of activity/functional limitations among persons with NTDs. Administration of the SALSA scale is easy, practical and is focused specifically on activity limitations contrary to the WHODAS (Nardi et al., 2012) which also focuses on participation restrictions. Data from validations using the P-scale Short in combination with either the WHODAS 2.0 or the SALSA should be compared to decide which of the two tools fits their aim best in order to contribute to a valid and useful generic cross-NTD toolkit. For further validation and use of the P-scale Short among other NTDs in differing environments, normative data should be collected to determine cut-off values of this shortened version. With regard to the cross-NTD character of the toolkit, it is important that additional validations recruit larger sample sizes and therefore could also focus on the cultural-equivalence of the tools included in the toolkit. Additionally, it is recommended to develop a user guide for the whole toolkit and finally to create an international database in which outcomes could be collected and monitored to identify global priority areas for MMD services.

REFERENCES

American Leprosy Missions by Fehman, L., Geyer, M. J. & Bolton, L. (2015). Ten Steps. A Guide for Health Promotion and Empowerment of People Affected by Neglected Tropical Diseases. *ALM.*

Bachani, A. M., Galiwango, E., Kadobera, D., Bentley, J. A., Bishai, D., Wegener, S., & Hyder, A. A. (2014). A new screening instrument for disability in low-income and middle-income settings: application at the Iganga-Mayuge Demographic Surveillance System (IM-DSS), Uganda. *BMJ open*, *4*(12), e005795.

Barry, J. A. (2014). Social sciences research on infectious diseases of poverty: too little and too late?

Berkowitz, A. L., Raibagkar, P., Pritt, B. S., & Mateen, F. J. (2015). Neurologic manifestations of the neglected tropical diseases. *Journal of the neurological sciences*, *349*(1), 20-32.

Brady, M. A., Hooper, P. J., & Ottesen, E. A. (2006). Projected benefits from integrating NTD programs in sub-Saharan Africa. *TRENDS in Parasitology*,*22*(7), 285-291.

Van Brakel, W. H., Anderson, A. M., Mutatkar, R. K., Bakirtzief, Z., Nicholls, P. G., Raju, M. S., & Das-Pattanayak, R. K. (2006). The Participation Scale: measuring a key concept in public health. *Disability & Rehabilitation*, *28*(4), 193-203.

Van Brakel, W. H., & Officer, A. L. A. N. A. (2008). Approaches and tools for measuring disability in low and middle-income countries. *Lepr Rev*, *79*(1), 50-64

Van Brakel, W. H., Sihombing, B., Djarir, H., Beise, K., Kusumawardhani, L., Yulihane, R., & Wilder-Smith, A. (2012). Disability in people affected by leprosy: the role of impairment, activity, social participation, stigma and discrimination. *Global health action*, *5*.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of cross-cultural psychology*, *1*(3), 185-216

Brislin, R. W. (1976). Comparative research methodology: cross-cultural studies. *International journal of psychology*, *11*(3), 215-229.

Brouwers, C., van Brakel, W., & Cornielje, H. (2011). Quality of life, perceived stigma, activity and participation of people with leprosy-related disabilities in south-east Nepal. *Disability, CBR & Inclusive Development*, *22*(1), 16-34.

Van Campen, C., Sixma, H. J., Kerssens, J. J., Peters, L., & Rasker, J. J. (1998). Assessing patients' priorities and perceptions of the quality of health care: the development of the QUOTE-Rheumatic-Patients instrument.*Rheumatology*, *37*(4), 362-368.

Cheung, M. K., Hung, A. T., Poon, P. K., Fong, D. Y., Li, L. S., Chow, E. S., ... & Liou, T. H. (2014). Validation of the World Health Organization Assessment Schedule II Chinese Traditional

Version (WHODAS II CT) in persons with disabilities and chronic illnesses for Chinese population. *Disability & Rehabilitation*, (0), 1-6.

Connaway, L. S., & Powell, R. R. (2010). *Basic research methods for librarians*. ABC-CLIO.

Cross, H., Arief, F., Beise, K., Brandsma, W., Chukwu, J. N., EBENSO, J., ... & LI, J. (2014). A Delphi exercise to refine the WHO three-point Disability Grading system for leprosy, and to develop guidelines to promote greater accuracy and reliability of WHO Disability recording. *Lepr Rev*, *85*, 18-28.

Davey, G. (2010). Podoconiosis, non-filarial elephantiasis, and lymphology. *Lymphology*, *43*(4), 168-177.

Ebenso, J., & Velema, J. P. (2009). Test-retest reliability of the screening activity limitation and safety awareness (SALSA) scale in North-West Nigeria. *Leprosy review*, *80*(2), 197.Eide, A. H. & Loeb, M. E. (2005) Data and statistics on disability in developing countries. *Disability KaR.*

Federici, S., Meloni, F., Mancini, A., Lauriola, M., & Olivetti Belardinelli, M. (2009). World Health Organisation disability assessment schedule II: contribution to the Italian validation. *Disability and rehabilitation*, *31*(7), 553-564.

Furtado, S. R., Sampaio, R. F., Vaz, D. V., Pinho, B. A., Nascimento, I. O., & Mancini, M. C. (2014). Brazilian version of the instrument of environmental assessment Craig Hospital Inventory of Environmental Factors (CHIEF): translation, cross-cultural adaptation and reliability. *Brazilian journal of physical therapy*, *18*(3), 259-267.

Grepin, K. A., & Reich, M. R. (2008). Conceptualizing integration: a framework for analysis applied to neglected tropical disease control partnerships.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field methods*, *18*(1), 59-82.

Han, C. W., Yajima, Y., Lee, E. J., Nakajima, K., Meguro, M., & Kohzuki, M. (2005). Validity and utility of the Craig Hospital Inventory of Environmental Factors for Korean community-dwelling elderly with or without stroke. *The Tohoku journal of experimental medicine*, *206*(1), 41-49.

Harrison-Felix, C. (2001). The Craig hospital inventory of environmental factors.*The Center for Outcome Measurement in Brain Injury*, 145-153.

Herdman, M., Fox-Rushby, J., & Badia, X. (1998). A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach.*Quality of life Research*, *7*(4), 323-335.

Hotez, P. J., Molyneux, D. H., Fenwick, A., Ottesen, E., Sachs, S. E., & Sachs, J. D. (2006). Incorporating a rapid-impact package for neglected tropical diseases with programs for HIV/AIDS, tuberculosis, and malaria. *PLoS medicine*, *3*(5), e102.

Hotez, P. J., Molyneux, D. H., Fenwick, A., Kumaresan, J., Sachs, S. E., Sachs, J. D., & Savioli, L. (2007). Control of neglected tropical diseases.*New England Journal of Medicine*, *357*(10), 1018-1027.

Hotez, P. J. (2008). Stigma: the stealth weapon of the NTD. *PLoS neglected tropical diseases*, *2*(4), e230.

Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation*, *12*(10), 1-8.

Ju, B., & Jin, T. (2013). Incorporating nonparametric statistics into Delphi studies in library and information science. *Information Research: An International Electronic Journal*, *18*(3), n3.

Kelders, R., van Brakel, W., & Beise, K. (2012). Testing and validating a simplified scale to measure social participation of people with disabilities in Indonesia. *Disability and rehabilitation*, *34*(8), 638-646.

Kersten, P., Cardol, M., George, S., Ward, C., Sibley, A., & White, B. (2007). Validity of the impact on participation and autonomy questionnaire: A comparison between two countries*. *Disability & Rehabilitation*, *29*(19), 1502-1509.

Kopparty, S. N. (1995). Problems, acceptance and social inequality: a study of the deformed leprosy patients and their families. *Leprosy review*, *66*(3), 239-249.

Lammie, P. J., Fenwick, A., & Utzinger, J. (2006). A blueprint for success: integration of neglected tropical disease control programmes. *Trends in parasitology*, *22*(7), 313-321.

Litt, E., Baker, M. C., & Molyneux, D. (2012). Neglected tropical diseases and mental health: a perspective on comorbidity. *Trends in parasitology*, *28*(5), 195-201.

Lunsford, T. R., & Lunsford, B. R. (1995). The Research Sample, Part I: Sampling. *JPO: Journal of Prosthetics and Orthotics*, *7*(3), 17A.

Miller, K., Mont, D., Maitland, A., Altman, B., & Madans, J. (2011). Results of a cross-national structured cognitive interviewing protocol to test measures of disability. *Quality & Quantity*, *45*(4), 801-815.

Mitchell, M., & Jolley, J. (2012). *Research design explained*. Cengage Learning.

Mont, D. (2007). Measuring disability prevalence. *World Bank Social Protection Discussion Paper*, *706*.

Nardi, S. M. T., Paschoal, V. D., & Zanetta, D. M. T. (2012). Limitations in activities of people affected by leprosy after completing multidrug therapy: application of the SALSA scale. *Leprosy review*, *83*(2), 172.

Parker, M. (1992). Re-assessing disability: the impact of schistosomal infection on daily activities among women in Gezira Province, Sudan. *Social science & medicine*, *35*(7), 877-890.

Parrillo, V. N., & Donoghue, C. (2005). Updating the Bogardus social distance studies: a new national survey. *The Social Science Journal*, *42*(2), 257-271.

Peters, R. M., Van Brakel, W. H., Zweekhorst, M. B., Damayanti, R., & Bunders, J. F. (2014). The Cultural Validation of Two Scales to Assess Social Stigma in Leprosy. *PLoS neglected tropical diseases*, *8*(11), e3274.

Rensen, C., Bandyopadhyay, S., Gopal, P. K., & Van Brakel, W. H. (2011). Measuring leprosy-related stigma-a pilot study to validate a toolkit of instruments. *Disability and Rehabilitation*, *33*(9), 711-719.

Riewpaiboon, A., Riewpaiboon, W., Ponsoongnern, K., & Van den Berg, B. (2009). Economic valuation of informal care in Asia: a case study of care for disabled stroke survivors in Thailand. *Social Science & Medicine*, *69*(4), 648-653.

Rouquette, A., & Falissard, B. (2011). Sample size requirements for the internal validation of psychiatric scales. *International journal of methods in psychiatric research*, *20*(4), 235-249.

Salsa Collaborative Study Group, & Salsa Collaborative Study Group. (2007). The development of a short questionnaire for screening of activity limitation and safety awareness (SALSA) in clients affected by leprosy or diabetes.*Disability and rehabilitation*, *29*(9), 689-700.

Saxena, R., Vashist, P., Singh, D., & Tandon, R. (2015). Preventing childhood blindness: Synergy between ophthalmology and community medicine. *Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine*, *40*(3), 149.

Silva, C., Coleta, I., Silva, A. G., Amaro, A., Alvarelhao, J., Queiros, A., & Rocha, N. (2013). Adaptation and validation of WHODAS 2.0 in patients with musculoskeletal pain. *Revista de Saúde Pública*, *47*(4), 752-758.

Silveira, C., Parpinelli, M. A., Pacagnella, R. C., de Camargo, R. S., Costa, M. L., Zanardi, D. M., ... & Andreucci, C. B. (2013). Adaptação transcultural da Escala de Avaliação de Incapacidades da Organização Mundial de Saúde (WHODAS 2.0) para o Português. *Revista da Associação Médica Brasileira*,*59*(3), 234-240.

Slim, F. J., van Schie, C. H., Keukenkamp, R., Faber, W. R., & Nollet, F. (2010). Effects of impairments on activities and participation in people affected by leprosy in The Netherlands. *Journal of rehabilitation medicine*, *42*(6), 536-543.

Stevelink, S. A. M., Van Brakel, W. H., & Augustine, V. (2011). Stigma and social participation in Southern India: Differences and commonalities among persons affected by leprosy and persons living with HIV/AIDS. *Psychology, health & medicine*, *16*(6), 695-707

Stevelink, S. A. M., Hoekstra, T., Nardi, S. M. T., van der Zee, C. H., Banstola, N., Premkumar, R., ... & van Brakel, W. H. (2012). Development and structural validation of a shortened version of the Participation Scale. *Disability and rehabilitation*, *34*(19), 1596-1607.

Stevelink, S. A. M., & van Brakel, W. H. (2013). The cross-cultural equivalence of participation instruments: a systematic review. *Disability and rehabilitation*,*35*(15), 1256-1268.

Suttiwong, J., Vongsirinavarat, M., Vachalathiti, R., & Chaiyawat, P. (2013). Impact on participation and autonomy questionnaire: psychometric properties of the Thai version. *Journal of physical therapy science*, *25*(7), 769.

Sudhir, P. M., Sharma, M. P., Mariamma, P., & Subbakrishna, D. K. (2012). Quality of life in anxiety disorders: Its relation to work and social functioning and dysfunctional cognitions: An exploratory study from India. *Asian journal of psychiatry*, *5*(4), 309-314.

Tazaki, M., Yamaguchi, T., Yatsunami, M., & Nakane, Y. (2014). Measuring functional health among the elderly: development of the Japanese version of the World Health Organization Disability Assessment Schedule II. *International Journal of Rehabilitation Research*, *37*(1), 48-53.

Tsutsumi, A., Izutsu, T., Islam, A. M., Maksuda, A. N., Kato, H., & Wakai, S. (2007). The quality of life, mental health, and perceived stigma of leprosy patients in Bangladesh. *Social science & medicine*, *64*(12), 2443-2453.

Üstün, B. (2001). Disability and culture: universalism and diversity.

Üstün, T. B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J. & Pull, C. (2010). Developing the World Health Organization disability assessment schedule 2.0. *Bulletin of the World Health Organization*,*88*(11), 815-823.

Üstün, T. B. (2010). *Measuring health and disability: manual for WHO disability assessment schedule WHODAS 2.0*. World Health Organization.

Weiss, M. G., Doongaji, D. R., Siddhartha, S., Wypij, D., Pathare, S., Bhatawdekar, M., & Fernandes, R. (1992). The Explanatory Model Interview Catalogue (EMIC). Contribution to cross-cultural research methods from a study of leprosy and mental health. *The British Journal of Psychiatry*, *160*(6), 819-830.

Weiss M. G., (2008). Stigma and the Social Burden of Neglected Tropical Diseases. *PLoS Negl Trop Dis* 2(5): e237.

Withington, S. G., Joha, S., Baird, D., Brink, M., & Brink, J. (2003). Assessing socio-economic factors in relation to stigmatization, impairment status, and selection for socio-economic rehabilitation: A 1 year cohort of new leprosy cases in north Bangladesh. *Leprosy review*, *74*(2), 120-132.

World Bank. (2015). Data. Retrieved May 6, 2015 from: http://data.worldbank.org/country/brazil

World health Organization. (2002). Toward a common language for function, disability, and health: ICF. World health Organization. Geneva. Retrieved February 25, 2015, from:

http://www.who.int/classifications/icf/icfbeginnersguide.pdf?ua=1

Geneva, W. H. O., & World Health Organization. (2012). Accelerating Work to Overcome the Global Impact of Neglected Tropical Diseases: a Roadmap for Implementation

World Health Organization. (2013 A). Sustaining the Drive to Overcome the Global Impact of Neglected Tropical Diseases. Second WHO Report on Neglected Tropical Diseases.

World Health Organization. (2013 B). How to use the ICF: A practical manual for using the International Classification of Functioning, Disability and Health (ICF). Exposure draft for comment.

World Health Organization. (2013 C). Lymphatic filariasis: managing morbidity and preventing disability: an aide-mémoire for national programme managers.Global Programme to Eliminate Lymphatic Filariasis.

World Health Organization (2015). Neglected tropical diseases. Fact sheets relating to NTD. World health Organization. Geneva. Retrieved May 6, 2015 from: http://www.who.int/neglected_diseases/mediacentre/factsheet/en/

de Zeeuw, J., Douwstra, M., Omansen, T. F., Sopoh, G. E., Johnson, C., Phillips, R. O., ... & Stienstra, Y. (2014). Psychometric properties of the participation scale among former Buruli ulcer patients in Ghana and Benin.*PLoS neglected tropical diseases*, *8*(11), e3254.

Zeldenryk, L., Gordon, S., Gray, M., Speare, R., & Melrose, W. (2012). Disability measurement for lymphatic filariasis: A review of generic tools used within morbidity management programs.

APPENDICES

Appendix 1. Cross-cultural equivalence checklist

I included my own version, adapted from Steveling & Van Brakel (2013), based on Bowden & Fox-Rushby (2003), Herdman et al. (1998), Terwee et al. (2007) and Mokkink et al. (2010).

Equivalence Category	Subcategories	Further Explanation	Assessment Values
Conceptual equivalence	C1. In what ways were the local populations' conceptualizations of participation assessed?	° Local literature ° Local questionnaires/instruments ° Discussion amongst researchers ° Involvement of anthropologists, sociologists, etc., ° Discussion with local people ° Other	'Positive' if 3 out of 6 subcategories received a positive rating.
	C2. Were any people of the target population asked to judge the appropriateness of the instrument; was a detailed discussion provided in the article concerning the appropriateness of the instrument or were the domains of importance identified by the local people covered in the instrument?		'Positive' if the answer is "yes".
	C3. Were any theoretical arguments presented questioning or accepting conceptual equivalence?	° Conceptual framework described in relation to the local concept under investigation ° Definition of the main construct ° Discussion of possible between-group differences related to construct ° Discussion of possible cultural differences related to the construct	'Positive' if 2 out of 4 subcategories received a positive rating.

Item equivalence	I1. Does the report mention how the authors assessed the relevance and acceptability of the individual items for the target population?		'Positive' if the answer is "yes".
	I2. Are the relevancy and acceptability of items discussed in the light of any quantitative or qualitative analyses?		
	I3. Were any adaptations necessary and was this discussed properly regarding individual items?		
Semantic equivalence	S1. Were the initial developers of the scale contacted and what was the nature of the contact?		'Positive' if the answer is "yes".
	S2. Was a translation protocol followed or a user manual including translation instructions?		
	S3. Were any details about the translation procedure provided?	° Description of the translators ° Was the translation procedure adequate? (translation and back translation, native speakers, with and without knowledge of the particular topic) ° Was the translation checked with the target population? ° Was the translation quality judged by experts or researchers?	'Positive' if at least 2 out of 4 subcategories received a positive rating.
	S4. Was the initial meaning of key words and phrases investigated and if yes, how was this done?		'Positive' if the answer is "yes".
	S5. Were there any problems or difficulties reported with the translation?		
Operational equivalence	O1. What was the percentage missing data and what action was taken if the percentage was too high (>25% per item)?		Give percentage:….% 'Positive' if % < 25 per item

	O2. Was the same administration format used?	° Was a description provided about the literacy rates or educational level of the target population? ° Was the suitability of the questionnaire format discussed? ° Was the appropriateness of the item format evaluated and discussed? ° Was the appropriateness of the response options evaluated and discussed? ° Were instructions for interviewers available?	‘Positive’ if at least 2 out of 5 subcategories received a positive rating.
	O3. Was the instrument pre-tested before use?		‘Positive’ if the answer is “yes”.
Measurement equivalence	M1. How was content validity addressed?	° Is the measurement aim of the instrument described? ° Is the target population described? ° Are the concepts described that the instrument intend to measure? ° Were the target population and researchers or experts involved during item selection and reduction? (Often not applicable during cross-cultural validation)	‘Positive’ if at least 2 out of 4 subcategories received a positive rating.
	M2. How was construct validity of the instrument assessed?	° Were hypotheses formulated a priori and was the expected magnitude range and direction of the expected association stated? ° Was factor analysis applied on an adequate sample size (at least seven times the number of items)?	‘Positive’ if 1 out of 2 subcategories received a positive rating.

	M3. Was test-retest reliability and agreement assessed?	° How was intra or inter-interviewer reliability assessed and were the results found adequate (Intraclass correlation coefficients (ICCs) ≥0.70 or weighted kappa ≥0.70)? ° Showed the scale adequate internal consistency (Cronbach's alpha at least 0.70)? ° Were adequate agreement measures provided (e.g. Smallest Detectable Change, Minimal Important Change)	'Positive' if 1 out of 2 subcategories received a positive rating.
	M4. Were any floor or ceiling effects tested (<15%)?		'Positive' if % floor or ceiling effects < 25.
	M5. How was interpretability assessed and were the results found adequate (at least means and standard deviations of four subgroups provided and/or a Minimally Important Change defined)?		'Positive' if assessment of interpretability was included with adequate results.
	M6. How was responsiveness assessed, were the methods applied adequate, as well as the results found?		'Positive' if responsiveness was assessed adequately.
	M7. Were any Item Response Theory (IRT) methods applied (e.g. Rasch analysis)?		'Positive' if the answer is "yes".

Appendix 2. Strategic plan for MMD

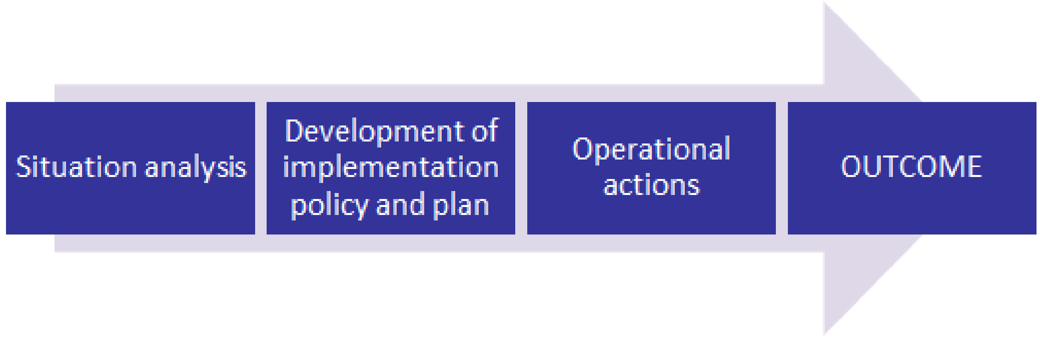


Figure A. Strategic plan for effective morbidity management and disability (MMD) (WHO, 2013 C).

In Figure A, the first step, ‘situation analysis’, includes mapping epidemiology of NTDs, health and social environment, and a strategic framework. This strategic framework is set up by analysed problems and gaps that are identified in the policy environment. Identification of these gaps could be done by similar activities as proposed by the WHO (2013 C) and listed in Table A. For each activity it should be indicated if there are policies already, guidelines or programmes and if or what the gaps need to be addressed.

Table A. Situation Analyses – Identification Framework

The proposed framework for identification of gaps in the policy environment in relation to morbidity management and disability of LF (MMD) (WHO, 2013 C), including translations of these activities to cross-NTD activities.

Activities(*Translation to cross-NTD activities)*
Treatment of acute dermatolymphangioadenitis. *(Treatment of acute NTD-related symptoms)*
Use of antibiotics *(*Use of pharmaceuticals)
Integrated control of NTDs
Water supply and sanitation improvement
Hygiene legalization
Surgery
Primary health care
Medical and clinical guidelines
Medical and clinical education
Community education and social mobilization
Psychosocial support
Socioeconomic rehabilitation
Collaboration with nongovernmental organizations
Other cross-NTD activities

The second step in the organization of MMD, ‘development of implementation policy and plan’, is carried out according to the situation analyses from the previous step. If policy is lacking in this situation analyses, the WHO suggests an example of content of a MMD implementation plan as can be found in Table B. This example is made by the WHO (2013 C) to give an idea of information and action that is needed to implement morbidity management and disability (MMD) if there is no policy yet. Implementation of MMD requires involvement of all stakeholders and clearly establishment of roles and responsibilities. Beside this implementers are asked to keep in mind advocacy and social mobilization according to the local circumstances of the implementation plan (WHO, 2013 C). The third and last step covers ‘operational actions’ which are classified according to two levels: national level and district and community level (WHO, 2013 C).

Table B. Example of Content of a MMD Implementation Plan

Country profile - Geography and climate - Political situation and administration - Demographic and socioeconomic information - Health status - Health system
Endemic situation and mass drug administration programme - Mapping - Microfilariae or antigenaemia baseline prevalence by implementation unit - Mass drug administration coverage - Clinical cases
Operational activities and responsibilities - Management and organization - Advocacy and social mobilization - Capacity building and training - Minimum packages of care for (acute) NTD-related symptoms - Psychosocial support and socioeconomic rehabilitation
Monitoring and evaluation - Data collection and analysis - Reporting
Time frame
Budget

Appendix 3. Phase 2: Systematic literature review

App. 3.1. Methods

App. 3.1.1. Purpose and organization

Considering the worldwide applicability within and across differing cultures of the generic cross-NTD toolkit, it was important to investigate known, cross-cultural validation studies of the relevant tools. Results of such an investigation could indicate the need for additional validations of a tool, before deciding whether to include it in the toolkit. Therefore a systematic literature search was done to collect countries and health-conditions for which the selected tools were already validated.

App. 3.1.2. Search strategy

A systematic literature review was conducted to investigate cross-cultural validation of tools that, according to the first results of the second e-survey from the Delphi study, had the most potential to be included in the toolkit and were not extensively validated already (such as many WHO instruments). Three databases, PubMed, Web of Knowledge and Science Direct were used to retrieve relevant publications using the search-terms given in App. 3.3.1. Filters on language were applied on those three databases (including only articles written in English). Additional publications that seemed relevant in the bibliography of-full-text articles were added through snowball selection.

Inclusion and exclusion criteria

English written, freely available and articles explicitly focused on cultural validation of one or more of the potential tools were included. Excluded were:

- articles mainly focusing on development of a tool,
- articles related to the selected tools but providing insufficient information about cultural validation of a tool (Stevelink & Van Brakel, 2013).

App 3.2. Results

The second phase of this study aimed to collect data concerning cross-cultural equivalence of the tools that have potential to be included in the generic cross-NTD toolkit by means of a systematic literature review.

Selected articles

Based on the tools mentioned by experts in the second e-survey, ten tools where included in the systematic literature review on their cross-cultural validation (Table D).

Table D. Tools included in the systematic literature review
The first column lists the tools which, during the second e-survey, received high total scores (second column) and were considered to be relevant for the systematic literature search.

Tool included in systematic literature review	Total points received in Delphi study
WHO (leprosy) disability grading system (with Eyes Hands Feet (EHF) as indicator)	12
WHO ICF checklist	35
Green Pastures Activity Scale (GPAS)	20
Washington Group Questionnaire on disability (WG questions)	29
Participation scale (P-scale)	*30*
Impact on Participation and Autonomy (IPA)	20
Craig Hospital Inventory of Environmental Factors (CHIEF) scale	5 participants promoted application 2 participants rejected application
Explanatory Model Interview Catalogue (EMIC) stigma scale - Individual	18
Explanatory Model Interview Catalogue (EMIC) stigma scale - Community	27
Social distance scale (SDS)	-

Not all highly ranked tools from the second e-survey were included in the systematic literature review. The WHODAS 2.0 and the SALSA for example both had high total scores (respectively 22 points and nineteen points), but were not included in the literature review. The main reason for exclusion of potential tools in the search for their cross cultural validity was that they already had been validated and/ or translated broadly in many languages as is the case with for example the WHODAS 2.0 which has been translated in at least 27 languages (Üstün, 2010) including validations in traditional Chinese (Cheung et al., 2014), Japanese (Tazaki et al., 2014), Italian (Ferderici et al., 2009) and European Portuguese (Silva et al., 2013). If such a broad validation was done, it was expected that there was no need for additional description on their cultural equivalence as an argument for their inclusion in the generic cross-NTD toolkit.

A total of 606 articles were retrieved from the systematic search on PubMed, Web of knowledge and Science direct. Figure 2 implies that Web of knowledge resulted in the highest number of articles (304 articles). Some articles found within the databases did overlap. After the initial search, seven out of the sixteen fully read articles were, based on the criteria as described in the methods, excluded after reading. Consequently, nine articles were included in this first level of the systematic literature review. These nine articles were checked on content and their bibliography was screened. This resulted in fourteen articles that were read to check their suitability. According to the in- and exclusion criteria, thirteen articles were not suitable and one was. The third level of the systematic literature review consisted of reading articles that were obtained from the bibliography from the article that was included in the second level. No new articles were included after checking this bibliography. In the end, ten articles on cultural validation of tools were included.

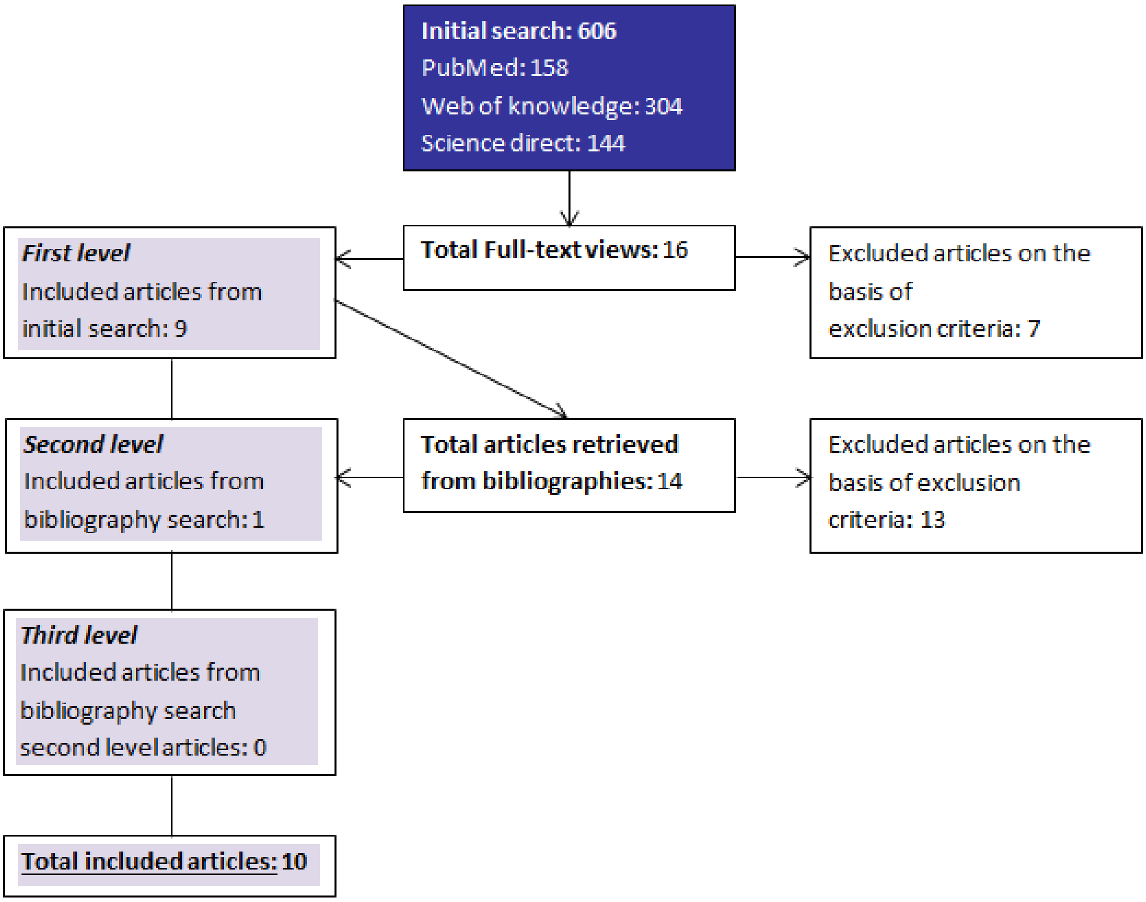


Figure 2. Flowchart displaying numbers on included and excluded of (full-text) articles within the systematic literature review
The initial search resulted in a total of 606 articles which were, after reading a total of 30 full-texts that seemed relevant from reading the abstracts and after excluding a total of twenty articles, reduced to ten articles.

Content of the included articles

App. 3.3.2. provides a detailed overview of the included articles divided over six tools (P-scale, IPA, CHIEF, EMIC individual and community, and SDS) in this literature study. No suitable articles were found for the WHO disability grading system with the EHF indicator, the WHO ICF checklist, the GPAS and the WG questions on disability.

Participant characteristics within the publications

Publications included study populations with communicable diseases (leprosy, polio, HIV/AIDS, tuberculosis and BU) and non-communicable diseases/ conditions (spinal cord injury, vitiligo, diabetes, stroke, neuromuscular disorder, rheumatoid arthritis, hand injury and cerebral palsy). Both the P-scale (Short) and the Social Distance Scale were validated among persons affected with leprosy. The EMIC individual and community were validated among persons affected with BU and leprosy.

Regions and languages in which tools are validated

Together, publications spanned ten countries divided over four of the six World Bank regions (http://data.worldbank.org/about/country-and-lending-groups, 11-5-2015). In table E it can be seen that most publications were about validations in East Asia & Pacific and in South Asia. Noticeably, four of the ten papers reported results from validation in India. Validations have been conducted in fourteen different languages of which most where Indian languages. The P-scale was validated most widely across cultures, while the IPA, the EMIC individual and the SDS were validated less widely. No publications were identified for Europe & Central Asia and for Middle East & North Africa.

Table E. Number of validation studies in each World Bank region per investigated tool

The first column lists the six tools that were included in the systematic literature search. The next six columns present the number of validations per World Bank region and in the last column total of World Bank regions are summed up.

Tool	Africa	East Asia & Pacific	Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	South Asia	Total number of regions
P-scale	*1*	*1*	-	*1*	-	*2*	***4***
IPA	-	1	-	-	-	-	**1**
CHIEF	-	1	-	1	-	-	**2**
EMIC individual	-	-	-	-	-	1	**1**
EMIC community	-	1	-	-	-	1	**2**
Social Distance Scale	-	1	-	-	-	-	**1**

Cross-cultural equivalence checking P-scale

For this study, only the retrieved cultural validation publications on the P-scale were assessed. From this assessment it became clear that internal consistency of the (adapted) P-scale within these studies was good, with a Cronbach's alpha ranging from 0.84 (Kelders et al., 2012) to 0.92 (Van Brakel et al., 2006). This means that there was a relatively high homogeneity among items within the P-scale. Another positive measurement result was that there were no floor and ceiling effects in application of the (adapted) P-scale (Van Brakel et al., 2006; Kelders et al, 2012; Jansen 2012; De Zeeuw et al., 2014). Adaptions made, were mainly shortening of the P-scale; Kelders et al. (2012) developed a simplified P-scale and Stevelink et al (2012) developed the 'P-scale Short'. Less to no information was found about the opinion of the target population regarding the tool and about missing data in the validation studies.

App 3.3. "App's"

App. 3.3.1. Search-Terms Systematic Literature Review

Craig Hospital Inventory of Environmental Factors. **CHIEF**

Pubmed: ("CHIEF"[Title/Abstract]) AND ("craig hospital inventory of environmental factors"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("CHIEF") AND TS=("craig hospital inventory of environmental factors") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("CHIEF") and TITLE-ABSTR-KEY("craig hospital inventory of environmental factors ") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

Explanatory Model Interview Catalogue. **EMIC stigma individual**

Pubmed: ("EMIC"[Title/Abstract]) OR ("explanatory model interview catalogue"[Title/Abstract]) AND (individu*[Text Word]) AND ("stigma"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("EMIC") AND TS=("explanatory model interview catalogue") AND TS=(individu*) AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("EMIC") or TITLE-ABSTR-KEY("explanatory model interview catalogue") and FULL-TEXT(individu*) and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-

KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

Explanatory Model Interview Catalogue. **EMIC stigma community**

Pubmed: ("EMIC"[Title/Abstract]) OR ("explanatory model interview catalogue"[Title/Abstract]) AND (community[Text Word]) AND ("stigma"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("EMIC") OR TS=("explanatory model interview catalogue") AND TS=(community) AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("EMIC") or TITLE-ABSTR-KEY("explanatory model interview catalogue") and FULL-TEXT(community*) and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

Green Pastures Activity Scale. **GPAS**

Pubmed: ("GPAS"[Title/Abstract]) OR ("Green Pastures Activity Scale" [Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("green pastures activity scale") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("Green pastures activity scale") or TITLE-ABSTR-KEY("GPAS") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence).

Impact on Participation and Autonomy. **IPA**

Pubmed: ("IPA"[Title/Abstract]) AND ("impact on participation and autonomy"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("IPA") AND TS=("impact on participation and autonomy") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("IPA") and TITLE-ABSTR-KEY("impact on participation and autonomy ") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

Participation scale. **P-scale**

Pubmed: ("P scale"[Title/Abstract]) OR ("participation scale" [Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("P scale") OR TS=("participation scale") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("P scale") or TITLE-ABSTR-KEY("participation scale") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

Social Distance Scale. **SDS**

Pubmed: ("SDS"[Title/Abstract]) AND ("social distance scale"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("SDS") AND TS=("social distance scale") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("SDS") and TITLE-ABSTR-KEY("social distance scale") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

Washington group questions on disability. **WG questions**

Pubmed: ("Washington group"[Title/Abstract]) AND (question*[Title/Abstract]) AND ("disability"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("Washington group") AND TS=(question*) AND TS=("disability") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("Washington group") and TITLE-ABSTR-KEY(question*) and TITLE-ABSTR-KEY("disability") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

WHO ICF checklist, impairments of body functions and structures

Pubmed: ("ICF" AND "checklist"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cu ltur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of knowledge: TS=("ICF" AND "checklist") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("ICF" and "checklist") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

WHO leprosy disability grading system (Eye-Hand-Foot score). **EHF-sore**

Pubmed: ("EHF”[Title/Abstract]) AND ("disability"[Title/Abstract]) AND (cross-cultur*[Title/Abstract] OR cultur*[Title/Abstract] OR valid*[Title/Abstract] OR equivalence[Title/Abstract])

Web of Knowledge: TS=("EHF") AND TS=("disability") AND TS=(cross-cultur* OR cultur* OR valid* OR equivalence)

Science Direct: TITLE-ABSTR-KEY("EHF") and TITLE-ABSTR-KEY("disability") and TITLE-ABSTR-KEY(cross-cultur*) or TITLE-ABSTR-KEY(cultur*) or TITLE-ABSTR-KEY(valid*) or TITLE-ABSTR-KEY(equivalence)

App. 3.3.2. Overview articles cross-cultural validation systematic literature search

Tool	[level found] Author	Country (language)	Study aim	NTD/ participant characteristics (N)	Adaptions/ application
P-scale	[1] Van Brakel et al. (2006)	India (Hindi, Bengali, Telugu and Tamil) Nepal (Nepali) Brazil (Potuguese)	"To develop a scale to measure (social) participation for use in rehabilitation, stigma reduction and social integration programmes."	Persons with leprosy, spinal cord injuries, polio and/or other disabilities. (initial N = 166, reduced to N = 30)	This was also a development study of the P-scale. The final, interview-based five-point response scale to assess the importance of the participation restriction (0=no restriction, 1=some restriction, but no problem, 2=small problem, 3=medium problem and 5=large problem) consisted of eighteen items.
	[1] Kelders et al. (2012)	Indonesia (not mentioned)	To test a simplified version of the P-scale in Indonesia against the original version, using the latter as the gold standard measure for participation.	Persons with different kinds of disability. Aged between 18 and 57, mean was 37 years. (N = 104; 47 males and 57 females)	Simplified version of the P-scale, interview. Three questions were omitted (No.5: "Do you take part in major festivals and rituals as your peers do? (e.g. weddings, funerals, religious festivals)"; No.12: "Do you move around inside and outside the house and around the village/neighborhood just as other people do?" and No.15: "In family discussions, does your opinion count?"). No.10 was replaced with the question '"Do you have the same opportunities as your peers to start or maintain a long-term relationship with a life partner?" and difficult words were also replaced.
	[other, not published yet] Jansen (2012)	India (Tamil)	"To validate the Participation Scale Short version (PSS) in Tamil	Persons with disabilities (PWD) due to leprosy, HIV/AIDS, tuberculosis, vitiligo	Here the P-scale Short (PSS) consisted of 14 questions after removing question No.7 "socially active", No.9

			Nadu, India, in order to assess the participation of people living with a disability. A related goal was to understand the concept of „participation‟ in this area.”	or diabetes. There was a qualitative study (PWD N = 10; 6 PWD and 4 professionals) and a quantitative study (N = 136; 86 PWD and 50 normative).	“caring for oneself”, No.15 “maintaining a relationship/marriage” No.16 “helping other people” and No.18 “learning new things”. Interview-based
	[1] De Zeeuw et al. (2014)	Ghana (Twi) Benin (French)	“To test the psychometric properties of the P-scale among former BU patients in Ghana and Benin.”	(former) Buruli ulcer patients (N = 143), their relatives (N = 137) and community controls (N = 106) (total N = 386).	Adapted for relatives and Items that not fitted the context of these countries were removed. Interview-based
IPA	[1] Kersten et al. (2007)	The Netherlands (Dutch) United Kingdom (English)	“To investigate the cross-cultural validity of the IPA.”	Persons known to rehabilitation services or GP practices with disabling conditions (stroke, neuromuscular disorder, severe hand injury, multiple sclerosis, rheumatoid arthritis, spinal cord injury or minor ailments). All participants were aged between 18 and 75. (N Dutch sample = 53; 36 males and 17 females and N English sample = 213; 89 males and 124 females)	No adaptions, the 31 participation and autonomy items were used.
	[1] Suttiwong et al. (2013)	Thailand (Thai)	“To cross-culturally translate and evaluate the reliability and validity of the Thai version of the Impact on Participation and Autonomy (IPA) in	Person with a traumatic spinal cord injury. Aged between 18 and 55. (N = 139; 110 males and 29 females)	No adaptions, only translation according to the guideline for cross-cultural adaptation of self-report measures.

			persons with spinal cord injury (SCI)."		
CHIEF	[1] Han et al. (2005)	Korea (Korean)	To develop the Korean version of the CHIEF and to evaluate its construct validity and utility.	Elderly aged 65 or above with or without stroke. (total N = 400; 109 males and 291 females)	Translation into Korean according to the Brislin's back translation technique (Brislin, 1970).
	[1] Furtado et al. (2014)	Brazil (Brazilian Portuguese)	"To translate the CHIEF questionnaire into Portuguese and cross-culturally adapt the questionnaire for the people of Brazil. Additionally, a second aim of this study was to determine the test-retest reliability of the questionnaire in a sample of caregivers of children and adolescents with cerebral palsy (CP)."	Caregivers of person with CP and aged between 24 and 59. (N = 47)	The term 'design and layout' was replaced by the Portuguese translation for 'physical structure'. Some other words were also changed. If needed, clarifying examples were provided verbally during the interview.
EMIC - Individual	[2] Weiss et al. (1992)	India (Hindi, Marathi)	Not mentioned	Person that just started leprosy treatment (N = 56) and a control group with skin diseases (N = 31).	Translation into Hindi and Marathi according to the Brislin's back translation technique (Brislin, 1976). Interview-based
EMIC - Community	[1] Rensen et al. (2011)	India (Tamill and Bengali)	"To compile a toolkit of stigma assessment instruments that can be used for baseline assessments and monitoring and evaluation of stigma reduction interventions."	Persons with leprosy (N = 806; 403 males and 403 females) and a control group (N = 165).	A 17 Likert item scale for leprosy-affected persons and a 13 Likert item scale for non-affected persons.
	[1] Peters et al.	Indonesia (Bahasa	To investigate the	Persons with Leprosy	The 15-item scale, adapted from the

	(2014)	Indonesia and if the respondent did not speak this language sufficient fluency, the questions were translated on the spot into Sudanese or Javanese.)	cultural validity of the Social distance scale and EMIC-community stigma scale in Cirebon District, Indonesia and additionally to bring a contribution to the body of knowledge regarding cross-cultural validation of instruments and assessment of social stigma in general.	(N = 259; 161 males and 98 females)	original EMIC.
Social distance scale	[1] Peters et al. (2014)	Same article as above.			

Appendix 4. First e-survey (Delphi study)

Introduction

Thank you for agreeing to participate in this survey on NTD-related morbidity and disabilities. The survey aims to make an inventory of the aspects of NTD-related morbidity and disability you and other NTD experts consider important for assessment and monitoring and of the tools used in the NTD field to assess these aspects. This information will be processed and, after one or more rounds of prioritization and selection, used to compile a generic cross-NTD toolkit.

This questionnaire is the first of two or three questionnaires included in this survey. In this first questionnaire more general topics are addressed. The topics of the next questionnaires will be developed based on the answers we receive in this first questionnaire. All questionnaires address NTD-related morbidity and disabilities and tools used in your field of expertise.

We expect to be able to send you the next questionnaire within a few weeks after this first round. Your help will be greatly appreciated.

Background

Various NTD-specific and generic tools have been in use to measure aspects of NTD-related morbidity and disability. The Participation Scale for example, was developed in patients affected by leprosy and other disabling conditions (van Brakel et al., 2006) but has now been used for other neglected tropical diseases as well. De Zeeuw and colleagues (2014) for example, used the Participation Scale to measure functional limitations and, in addition, used the Explanatory Model Interview Catalogue (EMIC) to measure perceived stigma in former Buruli ulcer patients in Ghana. Other general tools are the Hospital Anxiety Depression Scale (HAD), the Body Image Satisfaction Scale (BIS) and the Dermatology Quality of Life Scale (DQL) as Yanik and colleagues (2004) used to assess the psychological effect of cutaneous leishmaniasis. A more NTD-specific instrument was used by Harichandrakumar and colleagues (2006), who used the seven domains and five levels (7D5L) instrument to measure the health status of people with lymphatic filariasis in India.

Some NTDs, for example leprosy and lymphatic filariasis, are known to cause impairments such as mental health problems (Weiss, 2008) and stigma. Other domains in which disabilities may occur are activity / functioning and participation. All these aspects of functioning and health are included in the 'International Classification of Functioning, Disability and health' or 'ICF' (WHO, 2002). The ICF is an instrument to measure and classify health, functioning and disability in populations and individuals. It comprises the following interacting components: the health condition (NTD), body functions and structures, activity, participation, environmental factors and personal factors as can be seen in Figure 1. We will use this framework as a checklist for the components of morbidity and disability that may need to be assessed and monitor.

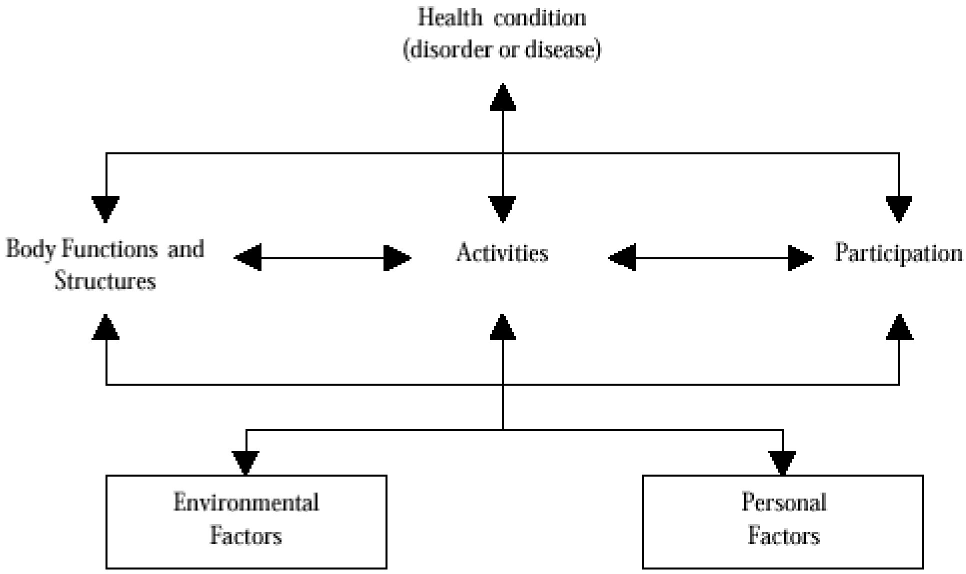


Figure 1: The ICF framework as developed by the WHO in 2001 (WHO, 2002).

More information about the ICF can be found with the following link:
http://www.who.int/classifications/icf/icfbeginnersguide.pdf?ua=1

References

de Zeeuw, J., Omansen, T. F., Douwstra, M., Barogui, Y. T., Agossadou, C., Sopoh, G. E., ... & Stientstra, Y. (2014). Persisting Social Participation Restrictions among Former Buruli Ulcer Patients in Ghana and Benin. *PLoS neglected tropical diseases*, *8*(11), e3303.

Harichandrakumar, K. T., Krishnamoorthy, K., Kumari, A. K., & Das, L. K. (2006). Health status of lymphatic filariasis assessed from patients using seven domains five levels (7D5L) instrument. *Acta tropica*, *99*(2), 137-143.

Tunde-Ayinmode, M. F., Akande, T. M., Ademola-Popola, D. S. (2011). Psychological and social adjustment to blindness: Understanding from teo groups of blind people in Ilorin, Nigeria. *Annals of African Medicine, 10, 155-163.*

Van Brakel, W. H., Anderson, A. M., Mutatkar, R. K., Bakirtzief, Z., Nicholls, P. G., Raju, M. S., & Das Pattanayak, R. K. (2006). The Participation Scale: measuring a key concept in public health. *Disability & Rehabilitation*, *28*(4), 193-203.

Weiss, M. G. (2008). Stigma and the social burden of neglected tropical diseases. *PLoS neglected tropical diseases*, 155-*2*(5), e237.

World Health Organization (2002). Toward a common language for function, disability, and health: ICF. World Health Organization, Geneva. Retrieved February 25, 2015, from *http://www.who.int/classifications/icf/icfbeginnersguide.pdf?ua=1*

Yanik, M., Gurel, M. S., Simsek, Z., & Kati, M. (2004). The psychological impact of cutaneous leishmaniasis. *Clinical and experimental dermatology*,*29*(5), 464-467.

Questionnaire

1) Would you please state your name and email address………………………………………………………………

What NTD(s) are your field of expertise?

Please select or more of the following NTDs:

- Buruli ulcer
- Chagas disease (American trypanosomiasis)
- Dengue
- Dracunculiasis (guinea-worm disease)
- Echinococcosis
- Endemic treponematoses (yaws)
- Foodborne trematode infections
- Human African trypanosomiasis (sleeping sickness)
- Leishmaniasis
- Leprosy (Hansen's disease)
- Lymphatic filariasis
- Onchocerciasis (river blindness)
- Podoconiosis
- Rabies
- Schistosomiasis
- Soil-transmitted helminthiasis (ascariasis, hookworm infection, trichuriasis)
- Teaniasis and (neuro)cysticercosis
- Trachoma
- Other (please specify):………………………………………………………………………………………………..

2) In what way do the NTDs you are working on cause chronic morbidity or disability?

3) What domains of the ICF conceptual framework are important in your opinion in measuring NTD-related morbidity and disability?

Choose one or more of the following domains:

- Health condition (disease complications)
- Body functions and structures (impairments)
- Activity (limitations; e.g. difficulties in learning, grasping objects, communication)
- Participation (restrictions; e.g. unemployment, no access to education, no social relationships)
- Environmental factors (e.g. health services, support, attitudes, products and technology)
- Personal factors (e.g. gender, age, socio-economic status)

4) What tools are used to assess NTD-related morbidity and disability in your field?

Please, briefly name the tools in the space below, indicating which domain of the ICF they assess:

5) Besides the domains listed in Q3, should any other domains be included in your opinion in measuring NTD-related morbidity and disability?

Domains listed in Q3:

- Health condition (disease complications)
- Body functions and structures (impairments)
- Activity (limitations; e.g. difficulties in learning, grasping objects, communication)
- Participation (restrictions; e.g. unemployment, no access to education, no social relationships)
- Environmental factors (e.g. health services, support, attitudes, products and technology)
- Personal factors (e.g. gender, age, socio-economic status)

If yes, please describe the domain(s) briefly in the space below:

6) Do you have any knowledge of tools that are used for other NTD(s) or disabilities that you believe could be important in assessment of different aspects of NTD-related morbidity and disability in your field?

If so, please name these tools briefly in the space below:

7) What do you think are the best interventions to facilitate solutions for morbidity management and disability prevention?

(E.g. a manual to organize focus groups with those who are affected and their environment)

Outro

Thank you for your contribution to this Delphi study, your help is greatly appreciated!
Would you please be so kind to send the tools that are used to assess NTD-related morbidity and disability in your field to ResearchNTD@Leprastichting.nl ? Thank you in advance. If you have any questions regarding this study, please do not hesitate to contact us.

Appendix 5. Second e-survey (Delphi study)

Introduction

Thank you for your willingness to participate in the Delphi panel supporting the development of a generic cross-NTD toolkit for assessment and monitoring of different aspects of NTD-related morbidity and disabilities. In addition to the Delphi study, we conducted a systematic literature review to make an inventory of the assessment/measurement tools currently used across the NTD field for morbidity and disability, as well as the tools that currently have been used to assess disability in general.

The toolkit is meant to be applicable for as many NTDs as possible. Tools should therefore preferably be generic. We are currently thinking of including the following domains: body functions and structures, activity, participation, environmental factors and personal factors.

This second questionnaire presents the tools that panel members have mentioned in the first questionnaire of this Delphi study (marked with an asterisk (*)). It also presents additional instruments found in the literature search we conducted and that may be of additional value for the toolkit. You are kindly asked to rank instruments for each domain according to priority as you perceive this. This way, we hope to be able to select the most appropriate and useful instruments to use across NTDs. You can also indicate when you consider an instrument to be essential for the NTD(s) you are working with, even though that instrument is NTD-specific. After all, we are compiling a toolkit and a toolkit may contain instruments for some very particular uses also. Your help and input is greatly appreciated!

Would you please state your name and email address:

1) Body functions and structures

Several instruments exist to assess body functions and structures. The first round of the Delphi study (marked with an asterisk (*)) and the literature search have so far given us the following instruments (see below). Could you please select the instrument or instruments that in your opinion have the best potential to use across NTDs:

1. WHO leprosy disability grading system* (with as indicator the Eye-Hand-Foot (EHF) score)
2. WHO ICF Checklist, impairments of body functions and structures*
3. The Buruli Ulcer patient's POD assessment form

In addition, we are thinking of including a tool to assess mental health. So far, the first round of the Delphi study and the literature search have given us the following instruments (see below). Could you please list, in order of preference, the three instruments that in your opinion have the best potential for use across NTDs:

1. Patient Health Questionnaire (PHQ-9)*
2. Self-Reporting Questionnaire (SRQ)*
3. Kessler Psychological Distress Scale (K10)*
4. Depression Anxiety Stress Scale-21 (DASS-21)*

1) Activity

Several instruments exist to assess activity or activity limitations. The first round of the Delphi study and the literature search have so far given us the following instruments (see below). Could you please list, in order of preference, the three instruments that in your opinion have the best potential for use across NTDs:

1. Screening of Activity Limitation and Safety Awareness (SALSA) scale*
2. World Health Organization Disability Assessment Schedule (WHODAS)*
3. Green Pastures Activity Scale (GPAS)
3. Barthel Index (BI)
5. Washington group questions on disability

2) Participation

Several instruments exist to assess participation restrictions. The first round of the Delphi study and the literature search have so far given us the following instruments (see below). Could you please list, in order of preference, the two instruments that in your opinion has the best potential for use across NTDs?

1. Participation scale (P-scale)*
2. Impact on Participation and Autonomy (IPA)
3. London Handicap Scale (LHS)
4. Sheehan Disability Scale (SDS)

3) Environmental factors

Several environmental factors can be assessed. The first round of the Delphi study and the literature search have so far given us the following factors that might be assessed

- Accessibility of services
- Accessibility of support
- Quality of services through the eyes of the patient
- Mapping of services in the community

- Economic support provisions (e.g. employment, vocational training)
- Facilitators and barriers
- Environmental hygiene

However, no specific tools were suggested for these. We suggest the Craig Hospital Inventory of Environmental Factors (CHIEF) scale as a tool to assess environmental factors. Can you give your opinion of the instrument in relation to the NTD(s) you work on? Do you know any other instruments that could be used across NTDs to assess environmental factors that should be assessed?

One important NTD-related factor is stigma. We will therefore include instruments to assess different aspects of stigma. Some of these relate to the environmental factors domain (community stigma and discrimination) and others to the personal factors domain). Several instruments exist to assess stigma. The first round of the Delphi study and the literature search have so far given us the following instruments (see below). Could you please select the instrument or instruments that in your opinion have the best potential for use across NTDs:

- Explanatory Model Interview Catalogue (EMIC) stigma scale community* – measures perceived stigma in the community
- Explanatory Model Interview Catalogue (EMIC) stigma scale individual* – measures perceived or anticipated stigma
- SARI Stigma scale* – measures anticipated stigma, disclosure concerns, internalized stigma and experienced stigma in four sub-scales
- Jacoby scale* - measures anticipated stigma
- Social Distance Scale* – measures attitude of the respondent towards condition described in a gender-specific vignette

4) Personal factors

Apart from age, gender, living area, religion, ethnic group, years of living with a certain condition, education, occupation, income and family structure - what other personal factors are in your opinion essential to assess? Please indicate if in your opinion some factors should not be recorded.

In addition to recording characteristics of participants, we would like to include a tool to assess motivation, empowerment or self-esteem, quality of life and perceptions/beliefs. The first round of the Delphi study and the literature search have so far given us the following instruments to assess personal factors. Could you please select what instruments, if any, you think should be assessed and included in the toolkit?

1. WHOQOL-BREF* - quality of life
2. Rosenberg Self-Esteem scale – self-esteem
3. I do not think a specific tool to assess these areas is necessary

5) Other

We strive to have a toolkit that would work across as many NTDs as possible, however, there may be a need to include disease-specific tools (such as the WHO disease-specific disability grading system; EHF score). In your opinion, do(es) the NTD(s) you work with need a disease-specific tool, and if so, for what domain of the ICF? Would you please list these areas below and include the disease-specific tool(s) required in your opinion

In your opinion, could these disease-specific tools be used or adapted for other NTDs as well, and if so, for which NTDs?

Thank you for your contribution to this round of the Delphi study, your help is greatly appreciated!
If you have any questions regarding this study, please do not hesitate to contact us at
.

Appendix 6. Third e-survey (Delphi study)

Introduction

Thank you for your participation in the Delphi study on the development of a generic cross-NTD toolkit for assessment and monitoring of different aspects of NTD-related morbidity and disabilities. This is the final round of the Delphi study.

This final round presents the proposed toolkit, based on tools that panel members have prioritised for each domain. This way, we tried to select the most appropriate and useful instruments to use across NTDs. In this round you will be asked to share your thoughts on the proposed toolkit. We are planning to do an initial validation across NTDs of some of the tools in the toolkit in an area where multiple NTDs are present in May and June 2015. In this round, you will also be asked to indicate which tools should be included in the initial validation of the toolkit.

Thank you once again for your participation. Your help and input is greatly appreciated!

Proposed composition of toolkit

Would you please state your name: *[typvlak]*

We are currently planning to include the following instruments in the toolkit:

1. WHO ICF Checklist - impairments of body functions and structures

2. Self-Reporting Questionnaire (SRQ)

3. World Health Organization Disability Assessment Schedule (WHODAS)

4. Participation scale (P-scale)

5. Explanatory Model Interview Catalogue (EMIC) - stigma scale affected persons

6. Craig Hospital Inventory of Environmental Factors (CHIEF) scale

7. WHOQOL-BREF with the WHOQOL-DIS (disability) module

Due to the limited time available, we are only able to include a selection of the tools in the initial validation. Could you please indicate, in order of priority, the instruments that in your opinion should definitely be included in the initial validation?
[typvlak]

Recently it has come to our attention that a WHO quality of life disability module (WHOQOL-DIS) exists to assess quality of life of people with disabilities. In previous Delphi rounds, the WHOQOL-BREF was proposed as a tool to assess quality of life. We have attached both the WHOQOL-BREF and the WHOQOL-DIS (which is an extended version of the WHOQOL-BREF) to the email you received. Which instrument would in your opinion be most suitable to be included in the toolkit?
[typvlak]

Additional remarks on the proposed toolkit, if any:
[typvlak]

Once the toolkit has been developed, you will receive a copy of the full report and the toolkit will be made available to the NTD community on the new information platform, InfoNTD. We plan to publish a summary of the results in a scientific paper in PLoS NTD. Would you be interested in being a co-author of this paper?
- Yes
- No

The toolkit will need several follow-up development phases during its development. Would you be willing to be contacted again as a resource person again during a follow-up phase?
- Yes
- No

Thank you
Thank you for your participation in our Delphi study. Your knowledge, time and help is greatly appreciated!

If you have any questions regarding this study, please do not hesitate to contact us at ResearchNTD@Leprastichting.nl.

Appendix 7. Clinical features

NTD	Clinical features
Chagas disease	• Indeterminate form • Cardiac form • Digestive form • Co-infection with HIV/aids (not obligatory)
Leishmaniasis	• Visceral leishmaniasis • Cutaneous leishmaniasis • Co-infection with HIV/aids (not obligatory)
Schistosomiasis	• Hepato-intestinal form • Hepatic form • Hepato-splenic form • Co-infection with HIV/aids (not obligatory)
Leprosy	• Indeterminate form • Tubercoloïde form • Borderline form • Lepromatous form • Co-infection with HIV/aids (not obligatory)

Appendix 8. Tool: WHODAS 2.0, 12-items

WHODAS 2.0

WORLD HEALTH ORGANIZATION
DISABILITY ASSESSMENT SCHEDULE 2.0

This questionnaire contains the interviewer-administered, 12-item version of WHODAS 2.0.

Instructions to the interviewer are written in bold and italics – do not read these aloud

Text for the respondent to hear is written in standard print in blue.. ***Read this text aloud***

Section 1 Face sheet

Complete items F1-F5 before starting each interview				
Item	**Question**			
F1.	Respondent identity number			
F2.	Interviewer identity number			
F3.	Assessment time point (1, 2, etc)			
F4.	Interview date	Day	Month	Year
F5	Living situation at time of interview (circle only one)			

Section 2 Demographic and background information

This interview has been developed by the World Health Organization (WHO) to better understand the difficulties people may have due to their health conditions. The information that you provide in this interview is confidential and will be used only for research. The interview will take 5–10 minutes to complete.

For respondents from the general population (not the clinical population) say:

Even if you are healthy and have no difficulties, I need to ask all of the questions so that the survey is complete.

I will start with some background questions.

Item	**Question**	**Answering options**	**Score**
A1	Record sex as observed	Female	1
		Male	2
A2.	How old are you now?	Years	
A3.	How many years in all did you spend studying in school, college or university?	Years	
A4.	What is your current marital status? (Select the single best option)	Never maried	1
		Currently married	2
		Seperated	3
		Divorced	4
		Widowed	5
		Cohabiting	6
A5.	Which describes your main work status best? (Select the single best option)	Paid work	1
		Self-employed, such as own your business or farming	2
		Non-paid work, such as volunteer or charity	3
		Student	4
		Keeping house/ homemaker	5
		Retired	6
		Unemployed (health reasons)	7
		Unemployed (other reasons)	8
		Other (specify)_____________ ____________________	9

Section 3 Preamble

Say to respondent:

The interview is about difficulties people have because of health conditions.

Hand flashcard #1 to respondent

By health condition I mean diseases or illnesses, or other health problems that may be short or long lasting; injuries; mental or emotional problems; and problems with alcohol or drugs.

Remember to keep all of your health problems in mind as you answer the questions. When I ask you about difficulties in doing an activity think about...

Point to flashcard #1

- Increased effort
- Discomfort or pain
- Slowness
- Changes in the way you do the activity.

When answering, I'd like you to think back over the past 30 days. I would also like you to answer these questions thinking about how much difficulty you have had, on average, over the past 30 days, while doing the activity as you usually do it.

Hand flashcard #2 to respondent

Use this scale when responding.

Read scale aloud:

None, mild, moderate, severe, extreme or cannot do.

Ensure that the respondent can easily see flashcards #1 and #2 throughout the interview.

Section 4 Core questions

Show flashcard #2

Item	In the past 30 days, how much difficulty did you have in:	None	Mild	Moderate	Severe	Extreme or cannot do
S1.	Standing for long periods such as 30 minutes?	1	2	3	4	5
S2.	Taking care of your household responsibilities?	1	2	3	4	5
S3.	Learning a new task, for example, learning how to get to a new place?	1	2	3	4	5
S4.	How much of a problem did you have joining in community activities (for example, festivities, religious or other activities) in the same way as anyone else can?	1	2	3	4	5
S5.	How much have you been emotionally affected by your health problems?	1	2	3	4	5

Item	In the past 30 days, how much difficulty did you have in:	None	Mild	Moderate	Severe	Extreme or cannot do
S6.	Concentrating on doing something for ten minutes?	1	2	3	4	5
S7.	Walking a long distance such as a kilometre [or equivalent]?	1	2	3	4	5
S8.	Washing your whole body?	1	2	3	4	5
S9.	Getting dressed?	1	2	3	4	5
S10.	Dealing with people you do not know?	1	2	3	4	5
S11.	Maintaining a friendship?	1	2	3	4	5
S12.	Your day-to-day work/school?	1	2	3	4	5

Item	Question	Number of days
H1.	Overall, in the past 30 days, how many days were these difficulties present?	
H2.	In the past 30 days, for how many days were you totally unable to carry out your usual activities or work because of any health condition?	
H3.	In the past 30 days, not counting the days that you were totally unable, for how many days did you cut back or reduce your usual activities or work because of any health condition?	

This concludes our interview. Thank you for participating.

Flashcard 1

Health conditions:

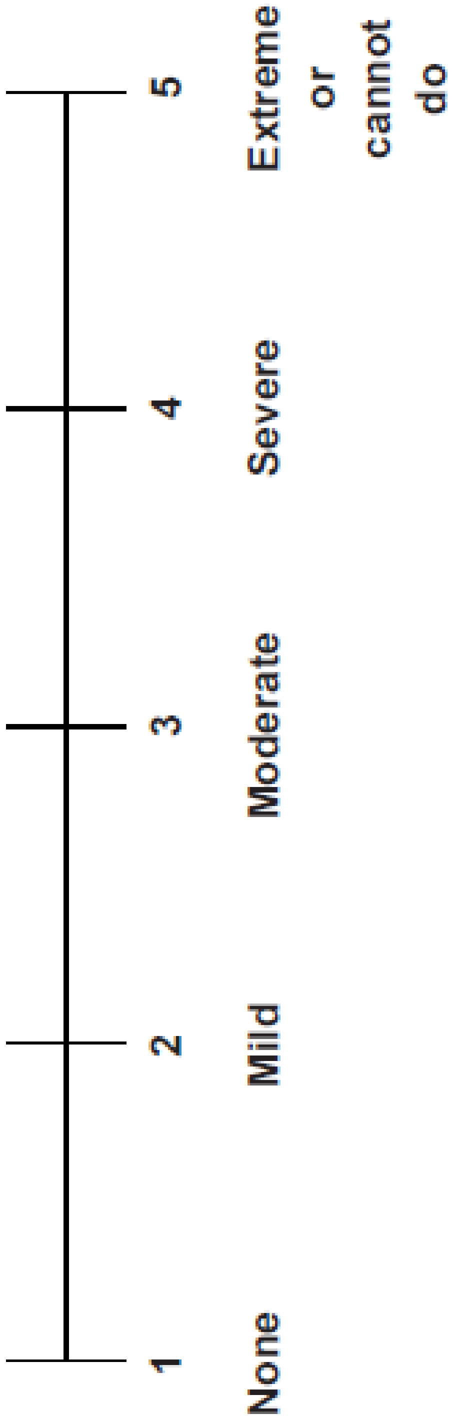
- **Diseases, illnesses or other health problems**
- **Injuries**
- **Mental or emotional problems**
- **Problems with alcohol**
- **Problems with drugs**

Having difficulty with an activity means:

- **Increased effort**
- **Discomfort or pain**
- **Slowness**
- **Changes in the way you do the activity**

Think about the past 30 days only.

Flashcard 2



Appendix 9. Tool: P-scale Short

No	Participation scale	Not specified, not answered	Yes	Sometimes	No	Irrelevant, I don't	NO problem	Small	Medium	Large	SCORE
1.	Do you have equal opportunity as your peers to find work?		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
2.	Do you work as hard as your peers do? (same hours, type of work etc)		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
3.	Do you contribute to the household economically in a similar way to your peers?		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
4.	Do you make visits outside your village / neighbourhood as much as your peers do? (except		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
5.	Do you take part in major festivals and rituals as your peers do? (e.g. weddings, funerals, religious		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
6.	Do you take part in social activities as much as your peers do? (e.g. sports, chat, meetings, religious or community		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
7.	Do you have the same respect in the community as your peers?		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
8.	Do you visit other people in the community as often as other people do?		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
9.	Do you move around inside and outside the house and around the village / neighbourhood just as other people		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
10.	In your village / neighbourhood, do you visit public places as often as other people do? (e.g. schools, shops,		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	

No	Participation scale	Not specified, not answered	Yes	Sometimes	No	Irrelevant, I don't	NO problem	Small	Medium	Large	SCORE
11.	In your home, do you do household work?		0								
	[if sometimes or no] How big a problem is it for you?						1	2	3	5	
12.	In family discussions, does your opinion count during important decision making?		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	
13.	Are you comfortable meeting new people?		0								
	[if sometimes or no] How big a problem is it to you?						1	2	3	5	