A comparative review of measurement instruments to inform and evaluate effectiveness of disability inclusive development

Nicolas Goujon¹, Alexandra Devine², Sally M. Baker², Beth Sprunt², Tanya J. Edmonds², Jennifer K. Booth², and Jill E. Keeffe¹

¹Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, University of Melbourne, East Melbourne, VIC, Australia and ²Nossal Institute for Global Health, The University of Melbourne, Carlton, VIC, Australia

Abstract

Purpose: A review of existing measurement instruments was conducted to examine their suitability to measure disability prevalence and assess quality of life, protection of disability rights and community participation by people with disabilities, specifically within the context of development programs in low and middle-income countries. Methods: From a search of PubMed and the grey literature, potentially relevant measurement instruments were identified and examined for their content and psychometric properties, where possible. Criteria for inclusion were: based on the WHO’s International Classification of Functioning Disability and Health (ICF), used quantitative methods, suitable for population-based studies of disability inclusive development in English and published after 1990. Characteristics of existing instruments were analysed according to components of the ICF and quality of life domains. Results: Ten instruments were identified and reviewed according to the criteria listed above. Each version of instruments was analysed separately. Only three instruments included a component on quality of life. Domains from the ICF that were addressed by some but not all instruments included the environment, technology and communication. Conclusion: The measurement instruments reviewed covered the range of elements required to measure disability-inclusion within development contexts. However no single measurement instrument has the capacity to measure both disability prevalence and changes in quality of life according to contemporary disability paradigms. The review of measurement instruments supports the need for developing an instrument specifically intended to measure disability inclusive practice within development programs.

Keywords

Development effectiveness, disability inclusive development, disability prevalence, evaluation, participation, quality of life, rights-based approach

Introduction

A new era commenced with the adoption of the United Nations Convention on the Rights of Persons with Disabilities (CRPD) in 2008 [1]. In accordance with Article 32, which promotes the right of people with disabilities to be included in all international cooperation efforts, increasing numbers of governments, non-governmental organisations and international funding agencies are moving towards the implementation of disability inclusive development [2–4]. Disability inclusive development consists of two main strategies: including the perspectives and rights of people with disabilities in all development activities, while at the same time empowering people with disabilities through disability specific projects [5].

It is estimated that people with disabilities represent approximately 15% of the world’s population and reside mainly in low and middle income countries [6]. Considered to be the world’s largest minority group [7], people with disabilities are more likely...
to live in poverty, [8] be illiterate [9] and experience social exclusion [10]. Not only are individuals with disabilities disadvantaged, but there is often a significant impact on the socio-economic situation of their families and the wider community [11]. With the increasing ageing population, and the emergence of non-communicable diseases in low and middle income countries [12], the number of people with impairments is predicted to further increase [6].

Despite these statistics, people with disabilities have been routinely excluded from development interventions which aim to reach the poorest of the poor. Article 32 of the CRPD specifies inclusion of people with disabilities within development activities, without which, there are serious concerns that the Millennium Development Goals – a unified set of development objectives addressing the needs of the world’s poorest and most marginalized people, intended to be achieved by 2015 – cannot be met [6,13].

There is a strong impetus, therefore, for development policies and programs which are inclusive of people with disabilities. However, this shift towards equity remains a challenge given the dearth of measurement instruments that support the design, implementation and evaluation of the effectiveness of disability inclusive activities. Currently no instruments exist that fulfill this requirement. An accurate measure of the prevalence of disability is crucial to the development of disability inclusive activities; so too is the ability to disaggregate data by type of impairment. Of equal importance is the capacity to assess the quality of life, awareness of rights and barriers to participation of people with disabilities in their community. To be considered effective, disability inclusive development programs should demonstrate improved participation of people with disabilities in their communities, such as improved access to education, health services and meaningful employment, contributing to an overall improvement in the quality of life of people with disabilities in low and middle income countries. To achieve this, programs firstly need to be able to identify people with disabilities, establish baseline measures on functioning, quality of life and access to services, and then be able to measure change in these areas post-intervention. The aim of this paper is therefore to review the ability of existing measurement instruments to contribute to these objectives; that is, identify people with disabilities, and measure whether development programmes improve their lives, using a rights-based paradigm.

**Defining disability: towards a rights-based approach**

No single definition of disability exists which is globally accepted and utilised. The concept of disability has evolved over recent decades. Distinct from the medical model, which promotes a view of people with disabilities as needing to be cared for or cured, there now exists several social models of disability. The Washington Group on Disability Statistics has developed small sets of disability measures for use in censuses, sample-based national or specialty surveys to provide information on the individual, socio-cultural perspectives. The global charter on disability rights, the CRPD, views disability through the lens of the rights-based approach. Two important elements of the rights-based approach are empowerment for the participation of people with disabilities and accountability [1]. The CRPD, defines disability as “an evolving concept... which results from the interaction between people with impairments and attitudinal and environmental barriers that hinder... full and effective participation in society on an equal basis with others”. Barriers can include those within the built environment, negative attitudes within communities or families, lack of access to information in appropriate formats, or policies which do not promote equal participation by people with disabilities.

Social determinants such as gender, age and ethnicity can also contribute to restricting or enabling a person’s empowerment in their community. It has been widely argued that women with impairment(s) can face multiple discrimination [7]. The capability approach developed by Sen [16], has moved the concept further, recognizing impairment as part of human diversity [17]. Within the capabilities approach, a person’s choices, expectations and empowerment are promoted as important factors for human development and achievements within their environment [14].

For the purpose of this study, disability has been defined in line with the CRPD and the ICF. This study assumes the right of people with disabilities to participate in society on an equal basis with others.

**Measuring disability**

Disability prevalence varies widely across and within countries. For example, Australia’s disability prevalence in 2003 was 20%, while that of the Cook Islands was reported to be 0.7% [18]. The Australian data come from the 2003 Australian Bureau of Statistics Survey on Disability, Ageing and Carers which used a question on limitations in daily activities in the previous 6 months while the Cook Islands data came from a register. It has been debated that methods to measure the prevalence of disability have lacked standardization [19]. This is partly a consequence of the shift in the disability paradigm, leading to various official definitions of disability. This is also due to use of different methodologies with varying efficacy to measure disability, e.g. variations in survey design and data collection methods such as self-report versus objective observation [20,21]. Thus, many countries lack valid and reliable disability information. With the exception of the UNICEF Multiple Indicator Cluster Surveys on child disability which have been undertaken in 18 countries using the UNICEF 10 questions instrument [22], data that do exist cannot be compared across countries and regions [20,21].

Therefore, countries generally under-report disability prevalence and the needs of people with disabilities [23]. This weakens the ability of development stakeholders to design appropriate and evidenced informed programmes, and measure the effectiveness of interventions which seek to target or include people with disabilities. The Washington Group on Disability Statistics has developed small sets of disability measures for use in censuses, sample-based national or specialty surveys to provide information on disability [24].

**Measuring quality of life**

Quality of life (QoL) is defined by the WHO Quality of Life Group as ‘‘individuals’ perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns’’ [16]. QoL can be seen as one outcome of the ICF domains, and one’s opportunities, empowerment and inclusion in one’s society. It has been suggested that QoL be added to models of disability [15] but it is not currently a part of the WHO Model of Disability in the ICF, even though the WHO definition of QoL is frequently cited. Two alternatives have been suggested to integrate QoL into the ICF conceptual scheme – either adding a separate construct to...
incorporate the generalized subjective well-being of QoL, or adding subjective components to each of the existing disability constructs. To intermingle items intended to measure both concepts in a single scale invites confusion and should be avoided, so it is recommended that both objective and subjective measures of participation are needed but as separate measures or as separate subscales in a single instrument [19].

Measuring change

Entities which implement disability-specific programmes typically use qualitative means to collect data to inform project design and assess their effectiveness. While this provides valuable information, it is difficult to quantify change. Disability-Adjusted Life Years (DALY) have been frequently used to assess the effect of public health interventions at the population level. DALYs were developed by the World Bank and the World Health Organization as a combined measure of morbidity and mortality, i.e. the sum of Years of Life Lost (YLL) and Years Lived with Disability (YLD) to compare burden of diseases. However, DALYs reflect the changing nature of medical conditions, and do not measure change in functional status, participation in society or Quality of Life thus not a suitable measure to measure change in individuals within a population [25]. While the ICF is composed of a comprehensive and consistent range of domains which encompass disability, it is too complex to be used itself either as a disability screening instrument. The ICF is also not designed to measure changes in quality of life of people with disabilities over time. In summary, constructs and issues relevant to evaluation of disability-inclusive development projects are: measuring disability prevalence, assessing change in QoL, assessing advancement of disability rights and access to the community.

Methods

To identify existing relevant measurement instruments, an initial literature search was conducted in January 2009 using the PubMed database (U.S. National library of Medicine, Bethesda, MD) and the Google™ browser (Google Inc., Mountain View, CA) and then continued till until July 2010 using PubMed alerts and the grey literature. Keywords are listed in Table 1. Only English language papers and reports published after 1990 were included in the review (Table 1). The search and review was conducted principally by NG with all co-authors contributing to the review of instruments and this paper. The team of authors and advisory group of three included people with extensive experience in the fields of disability, development and measurement tool design; most have worked in low resource countries in these fields. Consensus was reached within the team and advisory group on major decisions on methodology, inclusion of tools and the results of the review of tools.

Many scales have been developed to assess impairment and disability within the scope of the conceptual models presented, and to meet the needs of rehabilitation organisations [26]. The main inclusion criteria for this review included being based on the ICF domains and QoL assessment, and suitability for population-based research in developing countries (Table 1). From the search and after title and abstract review the list was reduced to approximately 30 instruments and finally 10 instruments were subjected to comparative review (Table 2). Measurement instruments were rejected because of the conceptual model of disability on which they are based (e.g. medical-based definition) [27], their cultural specificity [28] or because they were designed to be self-administered which excludes administration through a proxy for very young children and people with communication or literacy difficulties, who may require assistance [29].
In parallel to the ICF, the WHO developed the ICF checklist for research purposes. This instrument is compatible with the measurement of functional activity and participation domains, and environmental factors [34] (Table 3). This extensive checklist was not conceived as a ready-to-use questionnaire. The ICF checklist has been adapted and questionnaires designed for younger ages: <3 years, 3–6 years, 7–12 years and 13–17 years [34]. The length and format of these questionnaires with up to 90 items might present challenges for use in the field by development stakeholders. No information is available in the published literature about field testing of the checklist and questionnaires.

The WHO Disability Assessment Schedule II (WHO DAS II) [35] was designed to measure difficulties due to health conditions experienced by individuals or groups in the four weeks prior to administration of the questionnaire, and to monitor the impact of programmes through activity limitations and participation restriction domains described in the ICF. Two interview-based versions are available, which are comprised of 17 and 36 items (Table 3). The WHO DAS II has been assessed with groups of people with various health issues [30,36,37] and in different settings [38]. It has been also used as a reference instrument to measure disability prevalence [39] and to measure the outcomes of medical-related interventions [40]. A recent article provided details about the development process, field testing and psychometric properties of the instrument [41]. As stated by authors, the WHO DAS II focuses on the activities and participation domains of the ICF and does not measure the impact of environmental factors on life (Table 4). This limits its use to assess disability according to the CRPD and the ICF definitions, and to provide data critical for the design, implementation or evaluation of disability-inclusive development programmes.

The Activity Limitation Score and Participation Restriction Score (ALS/PRS) measurement of functional status, whilst based on the ICF differs from the ICF checklist in the areas of activities and participation by separating basic (e.g. walking) and complex activities (e.g. voting) into two independent scores (Table 3). The ALS/PRS has been developed from disability surveys in several African countries [42], and used alongside the WCG short set as a disability screening instrument [31]. It has been further tested for its ability to measure disability inclusiveness with a subsample of 1222 people from Zambia aged from 15 to 44 years identified as having a disability using the WCG short set [20].

### Table 2. Characteristics of the selected instruments.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Number of items</th>
<th>Target population (Years)</th>
<th>Type of response</th>
<th>Recall period (Months)</th>
<th>Comparison to other people</th>
<th>Using assistive device</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCG short set</td>
<td>6</td>
<td>&gt;5</td>
<td>RS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>WCG long set</td>
<td>33</td>
<td>&gt;5</td>
<td>RS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ICF Checklist/Questionnaire</td>
<td>30 to 90&lt;sup&gt;a&lt;/sup&gt;</td>
<td>All</td>
<td>RS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>WHO DAS II 12 Items</td>
<td>17&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&gt;18</td>
<td>RS</td>
<td>1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>WHO DAS II 36 Items</td>
<td>41&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&gt;18</td>
<td>RS</td>
<td>1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ALS/PRS</td>
<td>40</td>
<td>15–44&lt;sup&gt;c&lt;/sup&gt;</td>
<td>RS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>NSDA Screening Set</td>
<td>27</td>
<td>All</td>
<td>D</td>
<td>✓</td>
<td>5/27</td>
<td>✓</td>
</tr>
<tr>
<td>NSDA Health Set</td>
<td>19&lt;sup&gt;d&lt;/sup&gt; to 46</td>
<td>All</td>
<td>D + RS</td>
<td>6&lt;sup&gt;e&lt;/sup&gt;</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>NSDA Extensive Set</td>
<td>45&lt;sup&gt;f&lt;/sup&gt; to 120</td>
<td>All</td>
<td>RS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Participation Scale</td>
<td>18 to 36</td>
<td>NS</td>
<td>RS</td>
<td>✓</td>
<td>10/18 to 36</td>
<td>✓</td>
</tr>
<tr>
<td>WHOQOL-BREF</td>
<td>26</td>
<td>&gt;18&lt;sup&gt;g&lt;/sup&gt;</td>
<td>RS</td>
<td>1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CHIEF Short set</td>
<td>10 to 20</td>
<td>&gt;18</td>
<td>+</td>
<td>RS</td>
<td>12</td>
<td>✓</td>
</tr>
<tr>
<td>CHIEF Long set</td>
<td>25 to 50</td>
<td>&gt;18</td>
<td>RS</td>
<td>12</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>UNICEF 10Q</td>
<td>10</td>
<td>2–9</td>
<td>D</td>
<td>✓</td>
<td>3/10</td>
<td>✓</td>
</tr>
<tr>
<td>PedsQL&lt;sup&gt;TM&lt;/sup&gt; 4.0</td>
<td>23</td>
<td>2–18</td>
<td>RS</td>
<td>✓</td>
<td>1/23</td>
<td>✓</td>
</tr>
</tbody>
</table>

<sup>a</sup>Age related; 30 items to assess adult activities and participation' 90 items for assessment body functions, activities, participation and environment factors in children aged 7 to 12 years.

<sup>b</sup>Including health status assessment items added to the main core of 12 and 36 items.

<sup>c</sup>The sample was only people aged 15 to 44 years to analyse the validity of participation indicators (29).

<sup>d</sup>19 items for boys below 8 years (47), 46 items for people above 14 years.

<sup>e</sup>A recall period of 6 months is part of 5 of the 8 items for boys below 8 years (46) and 21 items of the 46 items for people above 14 years (47).

<sup>f</sup>45 items minimum for children below 12 years 120 items for adults (47).

<sup>g</sup>A survey conducted in Bangladesh showed promising results for the validation of the WHOQOL-BREF in the adolescent population studied (57).

### Legend for Tables 2, 3 and 4.

- **Absent/Not covered**: X
- **Present**: ✓
- **Partly represented/covered**: +
- **Largely mostly represented/covered**: ++
- **Comprehensively represented/covered**: ++++
- **Dichotomous**: +++++
- **Not Available**: NA
- **Not Applicable**: N/A
- **Not Specified**: NS
- **Response Scale**: RS

reference to the use of external assessment of respondent’s capacities, i.e. either their ability to execute a task or action (without assistance) or their performance in their current environment (with assistance) [15].

The WCG questionnaire on disability was developed as a universal and simple instrument to identify people aged five years and older who experience limitation(s) in activity performance (Table 2) [30]. A short set of 6 questions was specifically designed for census purposes and a long version with 27 additional questions on core domains was proposed and recommended to determine prevalence of limitations in functioning. The instrument has been tested in several countries including alongside validated measurement instruments [31] and census questions [32]. The WCG questions ask about impairment in line with functional ability to perform activities, but do not seek information about participation. Hence, according to the ICF definition of disability, WCG questions assess functional capacity, but not disability and do not measure all components of the ICF including major areas of life and attitudes (Tables 3 and 4). This limits its ability to measure disability and change in quality of life. The WCG is still undergoing a validation process [33] and is therefore not yet available for standardized screening of limitations in functioning.

In parallel to the ICF, the WHO developed the ICF checklist for research purposes. This instrument is compatible with the...
Table 3. Analysis of the selected instruments with the ICF model and Quality of Life d.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>ICF Checklist (adults)</th>
<th>Domestic life</th>
<th>Interpersonal interactions and relationships</th>
<th>Community social life</th>
<th>Mobility</th>
<th>Self-care</th>
<th>Communication</th>
<th>Body functions (impairments)</th>
<th>Learning and applying knowledge</th>
<th>Mobility</th>
<th>Self-care</th>
<th>Communication</th>
<th>Quality of life (QoL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCG short set</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WCG long set</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICF Checklist (adults)</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHO DAS 12 Items</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALS/PRS</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDSA Screening Set</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDSA Extensive Set</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHOQOL-BREF</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICEF Long set</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICEF Questionnaire (0–17 years)</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PedsQLTM 4.0c</td>
<td>+</td>
<td></td>
<td>+</td>
<td>×</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Social participation as defined by the ICF model and the capability approach proposed by Sen [16]. The general tasks and demands sub-category of the ICF taxonomy was discarded for its lack of specificity.

Quality of life (QoL) is evaluated through self-perception, awareness and final questions sections, mainly for people above 14 years (47).

Disability screening instrument as this instrument measures activity limitation rather than disability, and does not include items to identify people with mental illness. Environmental factors, of major importance for planning, implementing or evaluating disability-inclusive development activities are not specifically addressed in the ALS/PRS scale (Table 4). Within the development context, it is important to assess and measure change in access to the physical environment and also monitor environmental barriers such as attitudes within a community.

The National Disability Survey in Afghanistan (NDSA) was implemented to ‘assess disability prevalence, provide insight into the needs and opportunities of persons with disabilities, to identify environmental barriers and provide strategic guidelines to overcome the main difficulties faced by persons with disabilities in Afghanistan’ [42]. The authors developed an original instrument specific to the Afghan context. Its conceptual framework is the ICF model and the capability approach proposed by Sen [16]. A set of 27 questions is used as a screening instrument to assess a person’s impairment(s) and potential limitations in functioning. Due to contextual factors, the NDSA screening set and methodology is very specific to the Afghanistan context, and guidelines for its adaptation to other settings have not yet been developed; such guidelines could make it useful for other low income countries. The screening set has been designed for all ages, however reliance on a proxy to make judgements about young children can bring into question the validity instrument to identify all children (particularly those under five) with impairment who are at different stages of their development and thus to attribute. The NDSA Health set, composed of 9 modules and 46 questions evaluates deeper functioning abilities in regards to peoples’ daily life environment. A recall period of 6 months is used in five questions (Table 3). In addition, the authors developed extensive semi-structured sets (19 additional modules) on demographic characteristics, health status, education and work, livelihoods, social interactions, environmental factors and perception of disability (Tables 3 and 4) [43]. This extensive set has 120 items which limits its use by development organizations requiring measurement instruments which are easy-to-use [44]. To date, the NDSA survey design and field methodology have been thoroughly described, however, no published information about validation could be located [43].

The Participation Scale was designed to assess social inclusion with a set of 18–36 questions which address activities and participation domains defined in the ICF [45] (Table 3). It addresses most of the ICF activities and participation domains, with considerable emphasis on social activities (Table 4). It also includes an evaluation of discrimination from individuals, groups and societies (Table 4). Tested in three countries (Nepal, India and Brazil) with people with various physical impairments (sample size and distribution unknown), the scale was validated through self-appraisal of participants’ current situation on a visual analogue scale which was correlated with the participation scores. A moderate correlation was found with scores on the scale and experts rating the severity of participation restrictions. The use of external appraisal does not fit with the rights-based
approach which urges self-determination and participation. The use of comparison to peers or to other people is inconsistent in the Participation Scale and has been critiqued elsewhere [46]. The characteristics of the population assessed have not been reported and details of the psychometric properties were not available.

The WHO Quality of Life BREF (WHOQOL BREF) is a short version (30 items) of the WHOQOL-100, a QoL instrument [47]. The WHOQOL BREF assesses general QoL in the last four weeks prior to the testing. Inter- and intra-interviewer reliability were assessed demonstrating high/moderate agreement. It has been tested using classical validation methods such as discriminant ability and item response theory across ages, health conditions and countries [47–51]. The WHOQOL-BREF has been extensively used either to assess QoL [52,53], or as a reference instrument to validate the outcomes of interventions [54]. The WHOQOL-BREF explores some of the ICF domains, but is not designed to measure disability prevalence (Tables 3 and 4).

The Craig Hospital Inventory of Environmental Factors (CHIEF) was designed to determine how factors of the environment hinder or facilitate a person’s life achievements and can be used to explore how environmental factors act to contribute to disability [55]. Despite being specific to a single country, this was included as it is one of few instruments that include environmental factors. A set of 25–50 questions has been developed as well as a short form of 12–24 questions. Following wide consultations with advisory panels, the instrument was tested with two convenience samples from North America, a population-based sample of 2269 adults who self-rated their disability status in a telephone survey, and another with 409 adults with various types of impairments followed up in medical consultations. A recall period of 12 months is used (Table 3). The CHIEF scale showed very good reliability for test–retest but only moderate agreement for participant-proxy agreement. The total and five subscale scores show very good internal consistency. The instrument has shown good psychometric properties in discriminating between people with and without disability and between groups with different impairments but does not measure disability prevalence or quality of life (Table 4). More research is needed to confirm the ability of the instrument to measure environmental factors in population-based samples and in other socio-cultural settings.

**Measurement instruments developed for children**

The UNICEF ten questions (UNICEF 10Q) has been designed to measure limitations in functioning, i.e. impairments or health conditions, in children aged from 2 to 9 years in resource-poor settings and to screen children with a high risk of disability [20] (Table 3). It has been extensively used, mostly within the Multiple Indicators Cluster Survey (MICS) UNICEF programme. Due to its low-positive predictive value, a professional appraisal of disability is recommended [22,56]. This limits its use where professional appraisal is not available or appropriate. Tables 3 and 4 indicate that it assesses mostly body functions and some basic activities, rather than disability according to the ICF. Its use of language presents disability negatively, and not in line with the CRPD. An example is the item “Compared with other children of his/her age, does (name) appear in any way mentally backward, dull or slow?” The instrument is also not suitable for use with adults, adolescents or infants.

The PedsQL™ measurement model for Quality of Life Inventory TM Version 4.0 (PedsQL™ 4.0) is a scale developed to evaluate children’s health-related QoL using 23 items [57]. It measures four functioning domains: physical, emotional, social and educational. Designed originally for chronically ill children [58], multiple versions were developed for children and adolescents aged from 2 to 18 years (Table 3). Although extensively tested in regards of health conditions [59], no field testing has been reported in low income countries [60] or with children with disabilities [61]. The PedsQL™ 4.0 partly explores body functions, activity and participation domains of the ICF (Table 3).

**Discussion**

This review of existing instruments found that no single instrument encompasses the ICF in terms of domains and definition of disability, or has the capacity to be used to measure the effectiveness of disability inclusive development activities. Some measurement instruments, such as the WCG (short set) and the UNICEF 10Q focus more heavily on body functions and activity limitations, while not encompassing participation factors (Tables 3 and 4). Conversely, measurement instruments such as the long and short sets of the CHIEF assess the effects of the environment but not the personal components of the ICF. The child and adolescent versions of the ICF questionnaire address all domains of the ICF. However most measurement instruments, including the ICF adult questionnaire which capture some of the ICF components, do not include items to evaluate the effects of the human, institutional and built environments, all of which are important considerations to measure effectiveness of development. Whilst sensory and physical functions are usually included...
in all measurement instruments, mental health problems, which are experienced by up to 30% of the population in some developing countries [62], are not covered by the examined measurement instruments except for the WCG (long set) and the NDSA (screening set).

Parts of the instruments designed to assess perceived well-being [59,63], the WHODAS II and the NDSA (extensive set), also explore peoples’ satisfaction in life which emphasises the interaction between ICF domains and QoL [64]. However, if measures of participation and well-being are to be used to assess the effectiveness of development activities, the measurement instruments must provide quantitative results and have the ability to show change if it occurs. This means that the content of the required instrument needs to cover all relevant and important components and that the response scales need to be sensitive enough to detect change within a program cycle. Recent evaluation of QoL measurement instruments using methods based on item response theory have shown that many measurement instruments do not have a proven ability to make reliable comparisons before and after an intervention [65]. With the exception of the WHOQOL-BREF and WHO DAS II 36 items, none of the measurement instruments reviewed have been validated to demonstrate their sensitivity to detect change over time.

An effective prevalence instrument must be both sensitive and specific to only include people with impairment(s) and disability. The use of a recall period such as “in the last 6 months, have you...” can help to distinguish between a person with disabilities and a person with temporary decrease in functioning that does not necessarily result in disability. Neither the WCG nor the NDSA short set prevalence measurement instruments include a recall period but importantly both measurement instruments allow for a proxy to answer on behalf of a child or adult who cannot answer for themselves.

An instrument to gather prevalence data across all ages is needed. The WCG is suitable from age 5 years onwards. Specific issues related to infants are only partially considered by the ICF and adolescent questionnaires. The NDSA was designed to cover all ages however it is difficult to use for population-based prevalence as it does not fully consider developmental stages of infants and children. For instance, a child aged 13 months may be unable to walk and should not be considered as impaired whereas a 24-month-old child would be expected to do so. PEDS QOL and UNICEF 10Q are designed specifically as screening measurement instruments for children, so it cannot be used to obtain overall prevalence data. For children aged below 5 years, a reliable instrument to measure disability prevalence that considers developmental stages is urgently needed [66,67].

Research that establishes the reliability and validity of measurement instruments for prevalence studies is critical in the choice of an instrument to use in population-based studies and in development programs. Validation methods have been published for seven of the ten measurement instruments examined but not for the WCG sets, the ICF instruments and the NDSA questionnaire. Measurement instruments such as the ALS/PRS scale require further clarification of their scope, validation in different settings, or in establishment of their responsiveness to change in quality of life. Several measurement instruments require further testing in low and middle income countries to assess their application in such settings.

The availability of reliable data on the prevalence of disability both within and comparable between countries is essential to inform policy and programme development for disability inclusive development. Moreover, government, donors and non-government development implementers need information that can lead directly to action, i.e. implementation or adaptation of existing development programmes. With the exception of the CHIEF and the NDSA, the reviewed measurement instruments do not provide information relevant to development programs. The lack of a comprehensive measurement instrument applicable across different cultural contexts and easily used in the field reduces the development sector’s ability to measure the effectiveness of disability-inclusive projects.

Conclusion
This review demonstrates that while these measurement instruments have valuable qualities according to the context for which they were intended, no single instrument has the necessary characteristics to both measure disability prevalence and support the design, implementation and measurement of effectiveness of disability inclusive development programs. The authors have commenced a project that will draw on many aspects of the measurement instruments reviewed to develop an instrument which will facilitate measurement of disability prevalence, including barriers to participation experienced by people with disabilities, and their quality of life. This instrument will allow development implementers to inform program design through baseline studies and measure the effectiveness of disability inclusive development efforts.

Declaration of interest
The authors report no competing interests. No authors were paid by a pharmaceutical company or other agencies. AusAID had no role in the writing of the manuscript and submission for publication.

The research is funded by the Australian Agency for International Development (AusAID) through an Australian Development Research Award. CERA receives Operational Infrastructure Support from the Victorian Government.

Authors’ contributions
All authors made intellectual contribution to the writing of the manuscript.

NG performed literature search and designed figures. NG, AD, TJE, JKB and JEK interpreted the data. NG and JK wrote the paper. NG, AD, SMB, BS, THE, JKB and JEK edited the manuscript.

References


57. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. Med Care 2001;39:800–12.


