The AMC Linear Disability Score (ALDS): measuring disability in clinical studies
Weisscher, N.

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 2

The impact of disease-related impairments on disability and health-related quality of life: a systematic review

Nadine Weisscher, Rob J de Haan, Marinus Vermeulen

BMC Med Res Methodol. 2007;7:24
Abstract

Background: To investigate the interchangeability of measures of disability and health-related quality of life (HRQL) by comparing their associations patterns with disease-related impairment measures in patients with a variety of conditions.

Methods: A systematic literature search of MEDLINE, EMBASE, Web of Science and a hand search of reference lists through January 2006. Studies were included if they reported associations patterns between impairment and disability and between impairment and HRQL. Correlation coefficients were transformed to Fisher’s z effect size (ES(z)). Weighted averages were reported as pooled ES(z) with 95% confidence intervals (CI).

Results: The relationship between impairment and disability was stronger (pooled ES(z) = 0.69; 95% CI, 0.66 – 0.72) than between impairment and HRQL (pooled ES(z) = 0.38; 95% CI, 0.36 – 0.41). The physical component score (pooled ES(z) = 0.43; 95% CI, 0.39 – 0.47) and disease-specific HRQL (pooled ES(z) = 0.46; 95% CI, 0.40 – 0.51) were stronger associated with impairments than the mental component score (pooled ES(z) = 0.28; 95% CI, 0.20 – 0.36) and generic HRQL (pooled ES(z) = 0.36; 95% CI, 0.33 – 0.39).

Conclusions: This study shows measures of disability and different HRQL domains were not equally related to impairment. Patient’s impairments are better reflected in disability measures, than in HRQL instruments. There are many outcomes of interest and precisely defining them and measuring them will improve assessing the impact of new interventions.

Background

Choosing an outcome measure for use in clinical research is a complex process. If the outcomes are chosen inappropriately, a study may provide unreliable results. Frequently used outcomes measures are mortality rates, number of events (recurrent myocard infarction) or disease activity (lesion load on MRI).

Besides biological measures other levels of clinical measurement can be considered in clinical studies: impairments, disability and health-related quality of life (HRQL). Impairments are the direct organic manifestations of the disease such as consciousness and paresis. Disability can be defined as limitations in carrying out activities of daily living, such as self-care, mobility and activities inside or outside the home. HRQL refers to a broad spectrum of consequences of disease. Although this concept also includes elements of impairments and disabilities, it has a strong focus on patient’s social functioning and perceived health status and well being. Hence, impairment measures are closely related to the patient’s disorder, whereas the other outcome measures focus on the patient’s level of functional health. Each subsequent level of clinical measurement is increasingly less disease-specific and more relevant to the patient.

Recently, patient-relevant outcomes in terms of disability and HRQL have become more important. Although both functional outcomes reflect the consequences of diseases on personal level, they are conceptually different and not synonymous as is often thought by clinical researchers. Consequently, the decision to use one of these measures can have important implications for the interpretation of the study results. The objective of this systematic review is to investigate the interchangeability of measures of disability and HRQL by comparing their association patterns with disease-related impairment measures.

Methods

Data sources

We conducted a systematic literature search of MEDLINE from 1966 through January 2006, EMBASE from 1980 through January 2006 and Web of Science from 1988 through January 2006 to identify studies addressing the relation between outcome measures. A search strategy using Medical Subject Heading, text words and Publication Types Impairment or Body function (Publication date from 2001) or Body structure (Publication date from 2001) and Disability (Evaluation) or Activity or Disabled Persons or Activities of Daily Living and Health-related quality of life combined with Association or Evaluation Studies or Comparative Study or Validation Studies was used. Studies from English, German and French literature were included. The electronic search was supplemented by hand searching the investigators files, and retrieval of references cited in available literature. The search strategy was composed by one of the authors (NW) in consultation with a clinical librarian.

Study selection

A set of explicit criteria, composed by three investigators (NW, RdH, MV), were used for selection of the literature. Articles were included if they focused on methodological or metric aspects of patient-based outcomes (for example, methods of evaluating such measures, psychometric or clinimetric assessment of measures, comparative studies of measures).

Studies were included in which both the association between impairment and disability, and between impairment and HRQL were calculated by means of correlation coefficients or other association measures, regardless of the disease with the exception of psychiatric disorders. Hence, to improve the comparability between the correlation patterns between the different health concepts,
only studies were included in which the three types of health outcomes were assessed in the same patient population. The measurement instruments were questionnaires or observation lists. When a multi-scale questionnaire focuses on more than one concept we only included the sub-scale for the health domain in question. Excluded were studies which valued HRQL in terms of utilities or composite scores of conceptual different measures.

Data extraction
All data were independently abstracted by two investigators (NW, RdH) through use of a list of the criteria composed by three investigators. After study selection, measures were categorized according the definitions of the WHO International Classification of Impairments, Disabilities, and Handicap version \(^1\) and the revised International Classification of Functioning, Disability and Health.\(^3\) Outcome measures were classified into one of the following categories: impairment, disability, and HRQL. The latter was further categorized into generic HRQL, disease-specific HRQL, mental component of HRQL and physical component of HRQL. When there was doubt on the classification of the health domain a third reviewer (MV) was consulted.

A first draft synthesizing the data was produced by the first author of this review (NW) and extensively criticized by two other authors. The following information was abstracted: author, year of publication, sample size, disease, type of impairment and disability scales, distinction between generic or disease-specific and mental or physical HRQL, and the associations with impairment. Studies reporting repeated measures only baseline data was abstracted. Disagreements in abstracting were solved by discussion.

Statistical analysis
The correlations coefficients (\(r\)) published were converted using Fischer’s variance stabilizing z transform (ES(z)).\(^8\)

\[
ES(z) = \frac{1}{2} \log_e \left[ \frac{1 + r}{1 - r} \right]
\]

were the asymptotic variance was:

\[
v_{ES(z)} = \frac{1}{n-3}
\]

Mean ES(z) were calculated for studies reporting multiple correlations between the concepts studied.\(^8\) The ES(z) were combined using a fixed or random effects model, when appropriate. The weight for each study was inversely proportional to the square root of the number of patients included in the study. The methods for obtaining the pooled ES(z) and its 95% confidence interval (CI) are described in detail elsewhere.\(^11\) We labeled the strength of the ES 0.20 as small, 0.50 as medium and 0.80 as large.\(^11\) The pooled ES(z) were converted back to a correlation coefficient:\(^12\)

\[
r = \tanh(ES(z))
\]

in which we labeled the strength of the association; absolute values of \(r\), 0.00-0.19 is regarded as very weak, 0.20-0.39 as weak, 0.40-0.59 as moderate, 0.60-0.79 as strong and 0.80-1.00 as very strong correlation.\(^13\) The backward transformed correlation coefficients were also expressed in variance explained (\(r^2\)). The square of the correlation coefficient tells us what proportion of the variance of one variable can be explained by the linear correlation with the other.

Statistical heterogeneity was measured using the Cochrane Q statistics; \(p \leq 0.10\) was considered representative of significant statistical heterogeneity. To establish the effect of heterogeneity between studies on meta-analysis conclusions, we performed several subgroup analyses: impairment in relation to generic or disease-specific HRQL and impairment versus the physical or mental aspects of HRQL.

We constructed normal quantile plots to assess whether the ES(z) estimates are (approximately) normal distributed and to search for publication bias.\(^14\) With regard to the presence of publication bias we also used the Spearman’s rho rank correlation method\(^12\) to test the relationship between the standardized ES(z) and sample size (\(p < 0.05\) indicates significant publication bias) and a fail-safe number approach. A fail-safe number is the number of non-significant, unpublished, or missing studies that would need to be added to the meta-analysis in order to change the results of the analysis. If this number is relative large to the number of observed studies, the observed results, even with some publication bias, can be treated as a reliable estimate of the true effect.\(^15\) All analyses were performed in Meta-Win, version 2.0.

Results

Study characteristics
Our initial search yielded 434 potentially literature citations (Figure 1). Of these, 125 were excluded based on title (70 were investigating a psychiatric disorder, 44 were review articles, nine articles measured caregiver burden, one article reported data about a composite score and one reported HRQL in terms of utilities). Abstracts from 309 articles were retrieved and additional 188 articles were excluded (134 studies did not measure all three health domains, 49 did not report concrete associations between the health domains, three were investigating a psychiatric disorder, one reported data about a composite score and one reported HRQL in terms of utilities), leaving 121 articles for full review. Of these, 90 were excluded (44 did not report association patterns, 34 articles measured
not all three domains, in five articles impairment measures were no scale score, four articles used composite scores, one reported utility based HRQL, one concerned a proxy measure and one a double publication).

A total of 31 studies, with an average sample size of 401 (range 38 – 6497) persons, were found to conform to our inclusion criteria (Table 1). Seven studies reported data on outcome measures in patients with arthritis, six on Multiple Sclerosis, three on Parkinson’s disease, two on stroke, two on spinal cord injury, four on a variety of joint diseases (knee, patella, elbow, shoulder), one cervical dystonia, one diabetic neuropathy, one facial nerve paralysis, one AIDS, one chronic dysphonia, one chronic obstructive pulmonary disease, and one on psoriasis.

The studies included a wide range of disease-related impairment measures, as well as a variety of disability measures (Table 1), in which the Functional Independence Measure (FIM) was the most often used. With regard to the HRQL measures, 17 studies used the SF-36 (or SF-20), 12 studies reported both the physical component and mental component scores, in two studies the mental health dimension was considered as mental HRQL. Other scales that were considered as a mental component of HRQL were the psychosocial dimension of the Sickness Impact Profile (SIP), psychological well-being and the General Health Questionnaire (GHQ), the mental dimension of the Multiple Sclerosis HRQL scale (MSQL), the social well-being dimension of the Facial Disability Index (FDI), the Overall Rating of Life satisfaction (ORLS), and the emotional subscale of the Voice Handicap Index (VHI).

Overall the studies included 86 correlations between impairment and disability and 104 correlations between impairment and HRQL. Five studies reported only one single correlation, for the remaining studies mean ES(z) were calculated for each health domain. Table 2 presents effect sizes of the 31 studies included. Thirteen studies (42%) reported large ES(z) between impairment and disability (≥ 0.81), whereas 81% of the ES(z) related to HRQL was small to moderate (≤ 0.50). The ES(z) between impairment and the different types of HRQL were on average small to moderate, except for disease-specific HRQL; 50% of the studies reported moderate to large ES(z) (0.51 – 0.80).
## Table 1. Information on the 31 studies included in the analysis ordered by disease.

<table>
<thead>
<tr>
<th>References</th>
<th>N</th>
<th>Disease</th>
<th>Impairment</th>
<th>Disability</th>
<th>HRQL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Generic</td>
<td>Specific</td>
</tr>
<tr>
<td>Deyo, 1983</td>
<td>97</td>
<td>Arthritis</td>
<td>Morning stiffness, grip strength</td>
<td>SIP physical, ARA</td>
<td>SIP total, ARA</td>
</tr>
<tr>
<td>Birrell, 1996</td>
<td>86</td>
<td>Arthritis</td>
<td>OSRA</td>
<td>HAQ</td>
<td>MHAQ</td>
</tr>
<tr>
<td>Kien, 1996</td>
<td>1030</td>
<td>Arthritis</td>
<td>VAS pain, VAS fatigue</td>
<td>SIP</td>
<td>SIP</td>
</tr>
<tr>
<td>Wolfe, 2000</td>
<td>1773</td>
<td>Arthritis</td>
<td>Global GI severity</td>
<td>HAQ</td>
<td>HAQ</td>
</tr>
<tr>
<td>Angst, 2004a</td>
<td>43</td>
<td>Arthritis</td>
<td>cASES</td>
<td>pASES, DASH</td>
<td>SF-36</td>
</tr>
<tr>
<td>Angst, 2005a</td>
<td>79</td>
<td>Arthritis</td>
<td>cASES</td>
<td>pASES, DASH</td>
<td>SF-36</td>
</tr>
<tr>
<td>Salaffi, 2007</td>
<td>244</td>
<td>Arthritis</td>
<td>WOMAC-Pain, WOMAC-Stiffness Function</td>
<td>Function</td>
<td>SF-36</td>
</tr>
<tr>
<td>Rothwell, 1997</td>
<td>42</td>
<td>Multiple sclerosis</td>
<td>EDSS</td>
<td>OPCS</td>
<td>SF-36</td>
</tr>
<tr>
<td>Shamack, 1999</td>
<td>50</td>
<td>Multiple sclerosis</td>
<td>EDSS, SNRS</td>
<td>Barthel Index, FIM, EuroQol VAS, HQO</td>
<td>SF-36</td>
</tr>
<tr>
<td>Schwartz, 2001</td>
<td>274</td>
<td>Multiple sclerosis</td>
<td>SI</td>
<td>Disease steps, AI</td>
<td>SF-36</td>
</tr>
<tr>
<td>Miller, 2000</td>
<td>300</td>
<td>Multiple sclerosis</td>
<td>EDSS, MSFC</td>
<td>SIP physical</td>
<td>SF-36</td>
</tr>
<tr>
<td>Hobart, 2000</td>
<td>64</td>
<td>Multiple sclerosis</td>
<td>EDSS</td>
<td>Barthel Index, FIM, Well-being, GHQ</td>
<td>SF-36</td>
</tr>
<tr>
<td>Rasova, 2006</td>
<td>112</td>
<td>Multiple sclerosis</td>
<td>Muscle performance, Pulmonary ventilation</td>
<td>Barthel Index</td>
<td>MSQL</td>
</tr>
<tr>
<td>Rubenstein, 1998</td>
<td>193</td>
<td>Parkinson's disease</td>
<td>FSG (BADL, IADL)</td>
<td>SF-36</td>
<td>PDQ-39</td>
</tr>
<tr>
<td>Schrag, 2000</td>
<td>79</td>
<td>Parkinson's disease</td>
<td>HIX, UPDRS motor</td>
<td>EuroQol VAS, SF-36</td>
<td>PDQ-39</td>
</tr>
<tr>
<td>Franchignoni, 2005</td>
<td>70</td>
<td>Parkinson's disease</td>
<td>BBB, PCS</td>
<td>UPDRS-ADL</td>
<td>PDQ-39</td>
</tr>
<tr>
<td>de Haan, 1995</td>
<td>81</td>
<td>Stroke</td>
<td>Mathew, Orogogzo, SSS, NIHSS, CNS</td>
<td>Functional status, Tegner</td>
<td>SF-36</td>
</tr>
<tr>
<td>Gottlieb, 2001</td>
<td>100</td>
<td>Stroke</td>
<td>Class, Reding</td>
<td>FIM</td>
<td>LSI-a</td>
</tr>
<tr>
<td>Fuhrer, 1992</td>
<td>140</td>
<td>Spinal cord injury</td>
<td>ASIA</td>
<td>FIM</td>
<td>LSI-a</td>
</tr>
<tr>
<td>May, 2002</td>
<td>98</td>
<td>Spinal cord injury</td>
<td>ASIA</td>
<td>FIM</td>
<td>QLI</td>
</tr>
<tr>
<td>Marx, 2001</td>
<td>42</td>
<td>Krane disorders</td>
<td>AAOS, Lysholm</td>
<td>KAOS-ADL</td>
<td>SF-36</td>
</tr>
<tr>
<td>Padron, 2003</td>
<td>110</td>
<td>Patellar dislocation</td>
<td>Fulkerson, Kujala, LKDC, Lysholm</td>
<td>Tegner</td>
<td>MFA</td>
</tr>
<tr>
<td>MacDermid, 2001</td>
<td>70</td>
<td>Elbow pathology</td>
<td>ASES pain, FREE pain</td>
<td>DASH</td>
<td>SF-36</td>
</tr>
<tr>
<td>Beaton, 1990</td>
<td>90</td>
<td>Shoulder disorder</td>
<td>Elevation of the shoulder</td>
<td>SST, pASES</td>
<td>SF-36 (acute version) SF-20</td>
</tr>
<tr>
<td>Lindbeck, 1998</td>
<td>64</td>
<td>Cervical dystonia</td>
<td>TnI, TWSTRS-P</td>
<td>TWSTRS-D</td>
<td>QOL-ON PDQ-39</td>
</tr>
<tr>
<td>Vink, 2003</td>
<td>262</td>
<td>Diabetic neuropathy</td>
<td>TNS</td>
<td>QOL-ON-ADL</td>
<td>PDQ-39</td>
</tr>
<tr>
<td>Kahn, 2001</td>
<td>46</td>
<td>Facial nerve paralysis</td>
<td>FaceE</td>
<td>FDI – physical functioning</td>
<td>QOL-ON PDQ-39</td>
</tr>
<tr>
<td>Cleary, 1993</td>
<td>189</td>
<td>AIDS</td>
<td>Extreme pain, Total symptom score</td>
<td>FSG (BADL, IADL)</td>
<td>ORLS</td>
</tr>
<tr>
<td>Sporer, 2004</td>
<td>77</td>
<td>Chronic dysphonia</td>
<td>COPD</td>
<td>Voice disorder severity</td>
<td>Mental Health</td>
</tr>
<tr>
<td>Hu, 2006</td>
<td>58</td>
<td>Psoriasis</td>
<td>Breathlessness (VAS)</td>
<td>Impairment daily living</td>
<td>VHI</td>
</tr>
<tr>
<td>Zacharias, 2002</td>
<td>6497</td>
<td>Psoriasis</td>
<td>PASI, Severity</td>
<td>BDI, PSFDO, ADL</td>
<td>PLSI</td>
</tr>
</tbody>
</table>

### Abbreviations (in alphabetic order):
- AI: Ambulation Index.
- ARA: American Rheumatism Association’s functional classes.
- ASIA: Total Motor Index Score.
- (p)cASES: (patient) clinical American Shoulder and Elbow Surgeons questionnaire.
- BBS: Berg Balance Scale.
- BADL: basic activities of daily life.
- IADL: instrumental activities of daily life.
- BADL: basic activities of daily life.
- BDI: Baseline Dyspnea Index.
- CAMBS: Cambridge Multiple Sclerosis Basic Score.
- CNS: Canadian Neurological Scale.
- COPD: Chronic Obstructive Pulmonary Disease.
- DASH: Disability of the Arm, Shoulder and Hand questionnaire.
- EDSS: Kurtzke Expanded Disability Scale.
- FaCE: Facial Clinimetric Evaluation.
- FDI: Facial Disability Index.
- FIM: Functional Independence Measure.
- GHQ: General Health Questionnaire.
- GI: gastrointestinal severity.
- HAQ: Stanford Health Assessment Questionnaire.
- HSQ: Health Status Questionnaire.
- H&Y: Hoehn and Yahr.
- IADL: instrumental activities of daily life.
- KOS: Knee Outcome Scale.
- LSIa: Life Satisfaction Index.
- MFA: Musculoskeletal Functional Assessment.
- MSQL: Multiple Sclerosis Quality of life index.
- NIHSS: National Institutes of Health Stroke Scale.
- OPQS: Office of Population Census and Survey disability scale.
- OQLS: Overall Rating of Life Satisfaction.
- PASI: Psoriasis Area and Severity Index.
- PCSI: Postural Changes Scale.
- PD: Psoriasis Disability Index.
- PDQ-39: Parkinson Disease Questionnaire.
- PDI: Psoriasis Disability Index.
- PLSI: Psoriasis Life Stress Inventory.
- PMSS: Pain and Mobility Scale.
- QOL-ON: Norfolk Quality of Life Questionnaire-Diabetic Neuropathy.
- SF-20: Medical Outcome Study 20-Item Short-Form Health Survey.
- SF-36: Medical Outcome Study 36-Item.
- SSI: Symptom Inventory.
- SNRS: Scripps Neurological Rating Scale.
- SSS: Scandinavian Stroke Scale.
- SST: Simple Shoulder Test.
- TNS: Total Neuropathy Score.
- TWSTRS-D: Toronto Western Spasmodic Torticollis Rating Scale.
- VHI: Voice Handicap Index.
- WOMAC: Western Ontario and McMaster Universities.
Data synthesis

Review of the normal quantile plots (Figure 2 and 3) showed that one study\textsuperscript{46} substantially deviated from normality. Removal of this study revealed statistical homogeneity of the datasets for the association between impairment and disability (Q statistics, \(p = 0.34\)), and between impairment and HRQL (Q statistics, \(p = 0.15\)).

Table 2. ES(\(z\)) of the studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>References</th>
<th>generic (n = 23)</th>
<th>disease-specific (n = 10)</th>
<th>physical (n = 13)</th>
<th>mental (n = 21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deyo, 1983\textsuperscript{23}</td>
<td>0.23</td>
<td>0.31</td>
<td>0.65</td>
<td>0.32</td>
</tr>
<tr>
<td>Birrell, 1988\textsuperscript{24}</td>
<td>0.81</td>
<td>0.49</td>
<td>0.47</td>
<td>0.41</td>
</tr>
<tr>
<td>Kwen, 1998\textsuperscript{19}</td>
<td>0.50</td>
<td>0.39</td>
<td>0.29</td>
<td>-0.15</td>
</tr>
<tr>
<td>Wolfe, 2000\textsuperscript{26}</td>
<td>0.35</td>
<td>0.38</td>
<td>0.41</td>
<td>-0.17</td>
</tr>
<tr>
<td>Angst, 2004\textsuperscript{17}</td>
<td>0.60</td>
<td>0.11</td>
<td>0.41</td>
<td>0.51</td>
</tr>
<tr>
<td>Salaff, 2005\textsuperscript{28}</td>
<td>1.00</td>
<td>0.46</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Rothwell, 1997\textsuperscript{35}</td>
<td>1.30</td>
<td>0.15</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Sharrack, 1999\textsuperscript{36}</td>
<td>1.06</td>
<td>0.48</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>Schwartz, 1999\textsuperscript{27}</td>
<td>0.50</td>
<td>0.68</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>Miller, 2000\textsuperscript{38}</td>
<td>1.02</td>
<td>0.18</td>
<td>0.34</td>
<td>0.14</td>
</tr>
<tr>
<td>Hobart, 2000\textsuperscript{39}</td>
<td>1.33</td>
<td>0.06</td>
<td>0.34</td>
<td>0.14</td>
</tr>
<tr>
<td>Rasova, 2005\textsuperscript{40}</td>
<td>0.56</td>
<td>0.22</td>
<td>0.39</td>
<td>0.04</td>
</tr>
<tr>
<td>Rubenstein, 1998\textsuperscript{41}</td>
<td>0.73</td>
<td>0.40</td>
<td>0.48</td>
<td>0.31</td>
</tr>
<tr>
<td>Schrag, 2000\textsuperscript{42}</td>
<td>0.66</td>
<td>0.32</td>
<td>0.49</td>
<td>0.21</td>
</tr>
<tr>
<td>Franchignoni, 2005\textsuperscript{43}</td>
<td>0.79</td>
<td>0.71</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>de Haan, 1993\textsuperscript{44}</td>
<td>0.87</td>
<td>0.52</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Gottlieb, 2001\textsuperscript{45}</td>
<td>0.55</td>
<td>0.16</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Fuhrer, 1992\textsuperscript{46}</td>
<td>0.85</td>
<td>0.04</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>May, 2002\textsuperscript{47}</td>
<td>1.26</td>
<td>0.06</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Marx, 2001\textsuperscript{48}</td>
<td>1.19</td>
<td>0.35</td>
<td>0.83</td>
<td>0.01</td>
</tr>
<tr>
<td>Paxton, 2003\textsuperscript{49}</td>
<td>0.35</td>
<td>0.49</td>
<td>0.54</td>
<td>0.29</td>
</tr>
<tr>
<td>Maxer, 2001\textsuperscript{50}</td>
<td>0.85</td>
<td>0.41</td>
<td>0.54</td>
<td>0.29</td>
</tr>
<tr>
<td>Beaton, 1996\textsuperscript{51}</td>
<td>0.40</td>
<td>0.28</td>
<td>0.54</td>
<td>0.29</td>
</tr>
<tr>
<td>Lindeboom, 1998\textsuperscript{52}</td>
<td>0.23</td>
<td>0.21</td>
<td>0.54</td>
<td>0.29</td>
</tr>
<tr>
<td>Vink, 2005\textsuperscript{53}</td>
<td>0.66</td>
<td>0.59</td>
<td>0.54</td>
<td>0.29</td>
</tr>
<tr>
<td>Kahn, 2001\textsuperscript{54}</td>
<td>0.87</td>
<td>0.78</td>
<td>0.29</td>
<td>0.60</td>
</tr>
<tr>
<td>Olaya, 1995\textsuperscript{55}</td>
<td>0.60</td>
<td>0.38</td>
<td>0.29</td>
<td>0.60</td>
</tr>
<tr>
<td>Speyer, 2004\textsuperscript{56}</td>
<td>1.05</td>
<td>0.55</td>
<td>0.29</td>
<td>0.60</td>
</tr>
<tr>
<td>Hu, 2005\textsuperscript{57}</td>
<td>0.46</td>
<td>0.45</td>
<td>0.29</td>
<td>0.60</td>
</tr>
<tr>
<td>Zacharias, 2002\textsuperscript{58}</td>
<td>0.44</td>
<td>0.40</td>
<td>0.29</td>
<td>0.60</td>
</tr>
</tbody>
</table>

The plot compares two distributions by plotting their quantiles against each other. One of the distributions is the standard normal distribution (\(Y_{\text{as}}\)) while the other is the distribution one wishes to compare it against (here \(Y_{\text{as}} = \text{ES}(\text{z})_{\text{i}} / \text{v}_{\text{i}}\)). The deviations from linearity indicate deviations in normality. Publication bias will tend to leave strange gaps in the plot or lead to strange curves.
On aggregate level (Figure 4), there was a moderate to large association between impairment and disability (pooled ES(z) = 0.69; 95% CI, 0.66 – 0.72). The aggregated relationship between impairment and HRQL was statistically significant smaller (pooled ES(z) = 0.38; 95% CI, 0.36 – 0.41). Backward transformation in correlation coefficients resulted in a moderate association between impairment and disability (r = 0.59; r² = 36%) and a weak association between impairment and HRQL (r = 0.36; r² = 13%).

Subgroup analyses
Table 2 presents the ES(z) per study used for the subgroup analyses. Figure 4 also depicts the various impairment – HRQL subgroup analyses. No statistical heterogeneity could be demonstrated (range p-values of the Q statistics 0.35 – 0.44), with the exception of the association between impairment and the mental component of HRQL (p = 0.06).

The pooled ES(z) between impairment and generic HRQL measures was 0.36 (95% CI, 0.33 – 0.39; r = 0.35; r² = 12%).16-22,24-27,30,32,34-42,45 The pooled ES(z) between impairment and disease-specific HRQL measures was significantly higher (pooled ES(z) = 0.46; 95% CI, 0.40 – 0.51; r = 0.43; r² = 18%).16,21,29,30,33,35,36,40,43,44,45 Impairment measures were significantly higher related with the physical component of the HRQL measures (pooled ES(z) = 0.43; 95% CI, 0.39 – 0.47; r = 0.41; r² = 17%);16-17,19,20,26-31,33,35,36,38,39,40,43,44,45 than with the mental component (pooled ES(z) = 0.28; 95% CI, 0.20 – 0.36 [random effects model]; r = 0.27; r² = 7%).16,17,19,20,26-31,33,35,36,40,43,44,45

Publication bias
Review of normal quantile plots showed no clear indications for publication bias in both the analyses for the relationship impairment and disability and for impairment and HRQL (Figure 2 and 3). In addition, publication bias was not evident when the Spearman’s rho rank correlation method was used (both p > 0.50). The fail-safe calculation by Rosenthal’s method retrieved for both analyses a large number of studies (n = 27077 / n = 9755) needed to be added in order to change the results of the meta-analysis.
Discussion

In this study we investigated the impact of impairments on patients’ functional health in a spectrum of diseases. The results of the meta-analysis show that impairments explain 36% of the variance of the disability scores and 13% of the variance of HRQL scores. Impairment scales account for 17% of the physical dimension of HRQL, but can explain no more than 7% of the psychological aspects of HRQL.47

Due to the great variability in disease-related impairment scales and disability and HRQL instruments used in patient groups with a wide range of disorders, we expected heterogeneous data sets. Interestingly, after removal of one study which deviated from normality, homogeneity could be demonstrated in all analyses, except when we investigated the relationship between impairment and the mental component of HRQL. Possibly, this type of HRQL is less clearly defined compared to the other types of functional outcomes.

Treatments, such as new drugs or new surgical or radiological interventions, are aimed at reducing mortality and morbidity. If these treatments reduce impairments, the beneficial effects on functional health will be far better detectable in the patient’s level of disability than in his or her level of HRQL. Moreover, if the outcome is assessed in terms of HRQL, treatment effects will be more visible in the physical component than in the mental components, and will be better reflected in the disease-specific measures than in the generic HRQL measures. Therefore, outcome measures at the level of functioning of patients should be selected carefully when clinical studies are designed. It clearly does matter on which level the functional outcome measurement is chosen.

No doubt, HRQL scales represent an important measure from a patient’s point of view, but have serious disadvantages in many clinical efficacy studies. The extent to which patients fulfill social roles and participate in their environment cannot be observed directly and depends not only on their functional ability, but also on personal traits, social circumstances and societal barriers. In particular, HRQL scores, which reflect the patient’s perception of the consequences of disease, depend on numerous additional, usually psychosocial, factors other than the disease itself. This explains why the correlation between impairments and HRQL scores is weak, especially when emphasis is laid on subjective psychological scores of these scales. The subjective HRQL scores therefore are often beyond the influence of the physician. In the primary assessment of new drugs or new surgical or radiological procedures, there are many outcomes of interest and precisely defining them and measuring them will improve assessing the impact of new interventions. In contrast, disability (mobility, self-care) seem to be a more appropriate primary end point for assessing the functional health status of patients. They allow to perform basic and complex activities of daily life, as we showed, not only better related to the disease process itself, but is also observable and predictive of dependence,48 and forms an essential aspect of the patient’s HRQL,49 an aspect which is amenable for treatment.

A shortcoming of this meta-analysis might be that there are probably more studies than the 31 we found. Our search strategy was based on the ICIDH and ICF terms and few researchers distinguish between the terms impairment, disability and HRQL. If they distinguished between these terms correlations between impairment and disability or between impairment and HRQL were given rarely for all three domains together. Another point of discussion may be the concern for publication bias. However, the normal quantile plots, Spearman’s rank correlation coefficient method and fail-safe calculation method do not give an indication for this type of bias. Finally and unfortunately, we were not able to perform subgroup analysis based on the methodological quality of the studies. This is due to the psychometric and clinimetric articles analyzed used a hotchpotch of research designs which made it impossible to formulate unequivocal methodological criteria in terms of, for example, appropriateness of the design chosen, sample size consideration, blind assessment of outcome parameters, statistical adjustment for confounders, or loss to follow-up.50

Conclusions

New drugs or new surgical or radiological treatments are developed from a pathophysiologic perspective and are primarily directed at reducing impairments. Traditionally, biological and impairment measures are preferred. Increasingly, it becomes evident that this type of outcome is not always relevant for patients.21,51 Our study shows that patient’s impairments are clearly better reflected in disability measures, than in HRQL instruments. The relatively low association between impairment and HRQL is found in patients with various disorders, therefore we conclude that to assess the efficacy of a new treatment, disability is an important functional endpoint in clinical studies.

References

(7) Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. J Chronic Dis. 1978;31:70.


Chapter 2


Chapter 3

The AMC Linear Disability Score item bank: item response theory analysis in a mixed patient population

Rebecca Holman, Nadine Weisscher, Cees AW Glas, Marcel GW Dijkgraaf, Marinus Vermeulen, Rob J de Haan, Robert Lindeboom

Health Qual Life Outcomes. 2005;3:83