

Neutral Theory: Applicability and Neutrality of Using Generic Health-related Quality of Life Tools in Diseases or Conditions Where Specific Tools Are Available

Ravi Jandhyala (✉ ravi@medialis.co.uk)

Medialis Limited

Research Article

Keywords: SF-36, EQ-5D, utility, patient-reported outcomes (PRO), health-related quality of life, construct, Neutral Theory

Posted Date: December 28th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-132467/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 **Title: Neutral Theory: applicability and Neutrality of using generic health-**
2 **related quality of life tools in diseases or conditions where specific tools are**
3 **available**

4
5 **Author:** Ravi Jandhyala^{1,2}

6 ¹Medialis Limited, Banbury, UK

7 ²Pharmaceutical Medicine Research Centre, King's College London

8

9 **Corresponding Author**

10 Ravi Jandhyala

11 Email: ravi@medialis.co.uk

12 Telephone: +44 7717 448656

13 **ORCID iD: 0000-0002-7241-7476**

14

15 **Declarations**

16 **Ethics approval and consent to participate**

17 Not applicable

18

19 **Consent for publication**

20 Not applicable

21

22 **Availability of data and materials**

23 The datasets used and/or analysed during the current study are available from the
24 corresponding author on reasonable request.

25

1 **Competing interests**

2 The author declares that he has no competing interests.

3

4 **Funding**

5 No external funding sources supported this work.

6

7 **Author contributions**

8 RJ led and came up with the concept of the study and authored and approved the manuscript.

9

10 **Acknowledgements**

11 Jessica MacLeod (Medalis) provided support for the literature searches and evidence review,
12 sourcing the tools, and data extraction. Solomon Christopher (Medialis) conducted all the
13 statistical analyses.

14

15 **Authors' information**

16 RJ is a pharmaceutical physician specialising in medical affairs and is a visiting senior
17 lecturer at the Pharmaceutical Medicine Research Centre, King's College London. RJ is the
18 author of the Jandhyala method in observing consensus of answers to list-generating
19 questioning and has researched other methods of generating real-world evidence through
20 development of disease specific measurement tools and core datasets. He has a special
21 interest in the observation of the disease phenotype of subjects recruited to clinical studies.

22

23

24

1 **Abstract**

2 **Background:**

3 Health-related quality of life (HRQoL) tools are limited by the indicators included in the
4 construct and variation in interpretation by different researchers. Neutral Theory describes the
5 ideal construct that includes all relevant indicators and, therefore, complete accuracy, or
6 Neutrality. Neutral Theory can thereby provide the framework to develop or test constructs.
7 To assess the application of Neutral Theory, the Neutrality of generic tools (SF-36 and EQ-
8 5D) at measuring HRQoL was compared to disease/condition-specific tools, with the latter
9 considered surrogates for the Neutral construct.

10 **Methods:**

11 Full descriptions of all disease/condition-specific HRQoL tools published on PubMed (to 01-
12 Jul-19) were sourced. For each tool, the number of items with and without a direct match
13 within the SF-36 and EQ-5D was recorded and the sensitivity/specificity calculated.

14 **Results:**

15 The SF-36 and EQ-5D did not achieve a sensitivity/specificity both >50% against any of the
16 163 disease/condition-specific tools identified. At 20% prevalence of poor HRQoL, the false
17 positive rate (FPR) was >75% for all but two tools against the SF-36 and six tools against the
18 EQ-5D. Increasing poor HRQoL to 80%, 47 tools for the SF-36 and 48 tools for the EQ-5D
19 had a FPR <50%. For rare disease tools (<1/2,000 population; n=17), sensitivity/specificity
20 ranged from 0-40%/5-31% for the SF-36 and 0-22%/29-100% for the EQ-5D. For non-rare
21 (n=75) and symptom-specific tools (n=71) sensitivity/specificity was: 0-100%/0-100% (SF-
22 36) and 0-50%/0-100% (EQ-5D); and 0-60%/0-19% (SF-36) and 0-25%/0-100% (EQ-5D),
23 respectively. No concordance was recorded for 18% (2/11) of results from studies of rare
24 disease tools *versus* the SF-36 (no data *vs* EQ-5D). For non-rare, disease-specific tools,
25 results were discordant for 30% (25/84) and 35% (23/65) of studies against the SF-36 and

1 EQ-5D, respectively. For symptom-specific tools, corresponding results were 36% (24/66)
2 and 16% (5/31).

3 **Conclusions:**

4 Generic HRQoL tools appear poorly correlated with disease/condition-specific tools, which
5 indicates that adoption of Neutral Theory in the development and assessment of HRQoL tools
6 could improve their relevance, accuracy, and utility in economic evaluations of health
7 interventions.

8
9 **Key words:** SF-36; EQ-5D; utility; patient-reported outcomes (PRO); health-related quality
10 of life; construct; Neutral Theory

11

12

13

14

15

16

17

1 **Background**

2 Observation of a construct, such as a quality of life tool, first requires conceptualisation of the
3 construct at a theoretical level followed by its operationalisation at an empirical level.

4 Operationalisation involves selecting indicators to be measured in the observation of the
5 construct. Both are vulnerable to variation in interpretation by different researchers and can
6 result in a divergence in their measurement of the 'same' construct. A new theory recently
7 proposed is that of a 'Neutral Observer', which provides a framework on which a
8 determination of the Neutrality, or accuracy, of an observation of a given construct can be
9 based [1]. Neutral Theory represents the ideal and assumes a Neutral or exhaustive list of
10 relevant indicators in the construct, whereby the sensitivity and specificity are both 1 (*i.e.*,
11 100% accurate). The operationalisation of constructs using disease-specific indicators can
12 perhaps be considered closer to achieving Neutrality than those based on generic
13 observations.

14
15 Understanding the impact of treatment on patients' quality of life is a pivotal component in
16 the economic evaluation of health interventions. There is, however, no universally agreed
17 definition of the construct of quality of life, with the one provided by the World Health
18 Organization (WHO) perhaps the most commonly cited: "an individual's perception of their
19 position in life in the context of the culture and value systems in which they live and in
20 relation to their goals, expectations, standards and concerns" [2]. This broad definition
21 includes the person's physical health, psychological state, personal beliefs, social
22 relationships and their relationship to salient features of their environment. The WHO
23 definition, and other similar ones, were influential in the concept of health-related quality of
24 life (HRQoL), which refers to how well a person functions in their life and his or her
25 perceived wellbeing in physical, mental, and social domains of health [3].

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

Two frequently used tools that have been independently operationalised and aim to objectively assess HRQoL are: the Medical Outcomes Study Short Form family of measures (*e.g.* SF-36 [4, 5]) and the EuroQol five-dimensional (EQ-5D) [6, 7]. Both of these tools capture HRQoL across a series of domains or dimensions: vitality, physical functioning, bodily pain, general health perceptions, physical functioning, emotional functioning, social functioning, and mental health in the SF-36 [5]; and mobility, self-care, usual activities, pain/discomfort, and anxiety/depression in the EQ-5D [7]. Generic HRQoL tools have been widely adopted in Health Technology Assessments (HTAs), with the National Institute for Health and Care Excellence (NICE) in the UK recommending use of the EQ-5D in its Technology Appraisals [8]. While this allows for potentially more consistent, transparent and predictable decision-making, it is open to criticism, as generic measures may be insensitive or fail to capture important aspects of health for a specific disease or condition [9]. Disease- or condition-specific HRQoL tools have the advantage of being clinically relevant to the health problem and responsive to clinically important changes in state. Conversely, this specificity complicates comparisons with the general population and across treatments for different diseases, limiting their application in HTAs.

This study aimed to apply Neutral theory in assessing the Neutrality, or accuracy, and applicability of generic tools (SF-36 and EQ-5D) at measuring HRQoL in diseases or conditions where there is a specific tool available, to act a surrogate for the Neutral list in the measurement of HRQoL.

1 **Methods**

2 **Identification of disease- or condition-specific health-related quality of life tools**

3 A literature search was performed to identify all published disease- and condition-specific
4 HRQoL tools. Medline (PubMed) was searched through 01 July 2019 using the following
5 terms: [“patient reported outcome” OR “PRO” OR “Quality of life” OR “QoL” AND
6 “disease specific” OR “condition specific”]; limit: [English language]. Two reviewers
7 undertook the search, with initial screening of abstracts and titles conducted using the semi-
8 automated Rayyan tool (<https://rayyan.qcri.org/>) [10]. Full descriptions of the identified
9 disease/condition-specific HRQoL tools were sourced as were the SF-36 and EQ-5D. In
10 addition, all original studies where HRQoL was assessed using a disease/condition-specific
11 HRQoL tool and the SF-36 and/or the EQ-5D were reviewed.

12

13 **Inclusion of appropriate domains and items**

14 The risk that the generic tools (SF-36 and EQ-5D) might include irrelevant domains or items
15 or exclude relevant domains or items for a specific disease or condition was assessed. Firstly,
16 for each condition- or disease-specific tool the number of items with and without a direct
17 match to the SF-36 and EQ-5D was recorded (for the EQ-5D, it was permitted for each of the
18 five questions to cover more than one item in each disease/condition-specific tool). The
19 sensitivity and specificity of the generic tool *versus* the disease/condition-specific tool was
20 then calculated as follows. True positives represented items captured in both the
21 disease/condition specific and generic tool; false positives, those captured in the generic tool,
22 but not in the disease/condition specific tool; and false negatives, those captured in the
23 disease/condition specific tool, but not in the generic tool. Since it is not possible to know if
24 the disease/condition specific tool fully captures all relevant items or domains, the true
25 negative fraction was assumed to be 0.9 (*i.e.* an arbitrary 10% missing).

1 Sensitivity/specificity results were further stratified into rare diseases (defined as affecting <1
2 in 2,000 population) [11], non-rare diseases (≥ 1 in 2,000 population), and symptom-specific
3 tools (*i.e.* those that cover symptoms [*e.g.* anxiety and depression; respiratory problems *etc*]
4 that might be present in multiple diseases/conditions).

5

6 The potential for misclassification of patients' HRQoL by a generic tool was expressed as the
7 median proportion of false positives and false negatives (with 95% confidence intervals
8 [CIs]), based on 1,000 studies, with prevalence of poor HRQoL set at 20%, 50%, and 80%.

9

10 **Concordance of quality of life scores**

11 For each of the studies comparing a disease/condition-specific tool with the SF-36 and/or
12 EQ-5D, a measure of concordance of the results was assigned. *No (none)* concordance was
13 assigned if a significant impact on HRQoL was seen with the disease/condition specific tool,
14 but no change or the opposite impact was seen with the generic tool (or *vice versa*); *Moderate*
15 concordance if HRQoL impact was scored in the same direction with both tools, but was
16 statistically significant with only one of them; and *Strong* concordance if the results were
17 fully aligned (significant/non-significant impact in same direction). For studies that measured
18 HRQoL changes over time, it was determined whether the concordance between the generic
19 and disease/condition-specific tool varied at different time points. Results were split into rare
20 diseases non-rare diseases, and symptom-specific tools.

21

22 All analyses were performed using R 3.6.0 (Revolutions Analytics) and Microsoft Excel 365
23 (Microsoft).

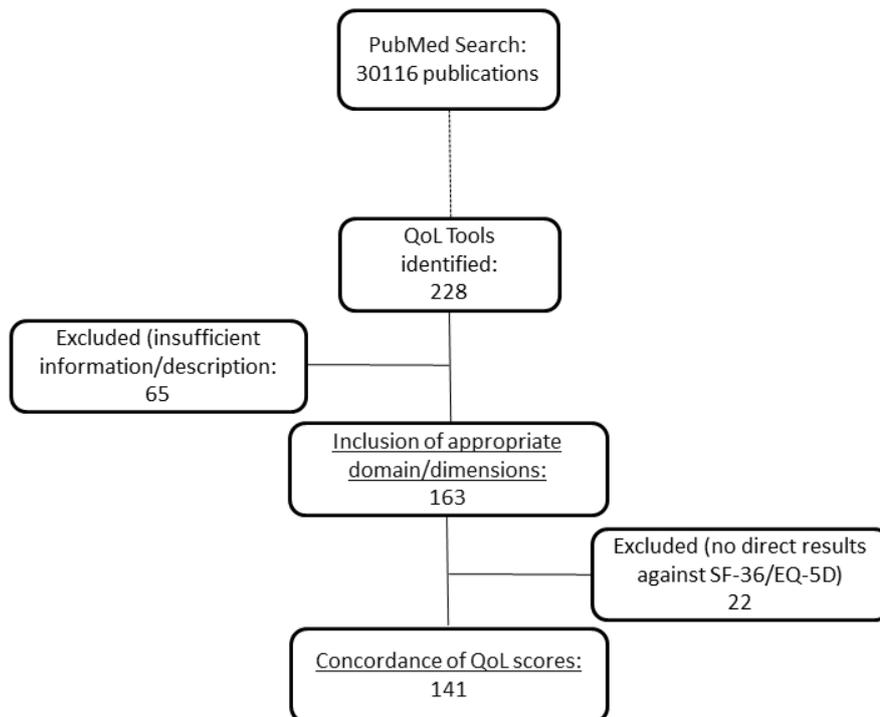
24 **Results**

25 **Identification of disease- or condition-specific quality of life tools**

1 A total of 30,116 publications were reviewed from which 228 disease- or condition-
2 specific HRQoL tools were identified (Figure 1 & Additional file 1). Full descriptions of 65
3 tools were unable to be sourced, either from published papers, online repositories, or *via*
4 direct approaches to the authors. The remaining 163 tools (rare diseases: 17; non-rare: 75;
5 symptom specific: 71) provided sufficient information/data for analysis, including 141
6 reporting results for a direct comparison against the SF-36 and/or EQ-5D (rare diseases: 10;
7 non-rare: 73; symptom specific: 58). One tool (University of California Los Angeles-Prostate
8 Cancer Index [UCLA-PCI]) completely overlaps with the SF-36, so was excluded from the
9 SF-36 comparisons.

10

11 **Fig. 1 Overview of identification and selection of disease- and condition-specific health-**
12 **related quality of life tools**



13

14

1 **Inclusion of appropriate domains and items**

2 *SF-36*

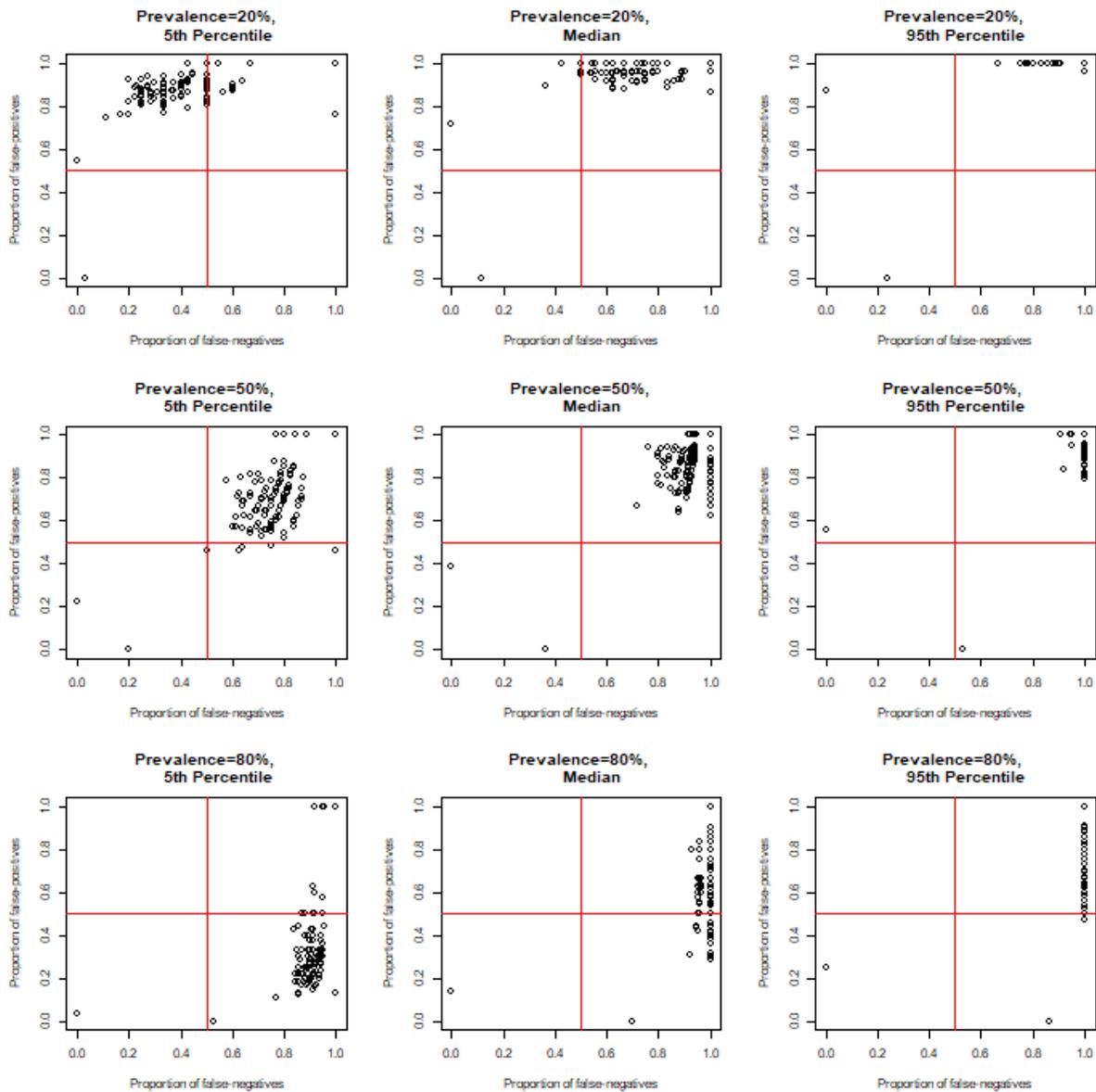
3 The SF-36 had a sensitivity of >75% against only one of the 162 disease/condition-specific
4 HRQoL tools analysed, with a further five tools having a sensitivity between 50-75%
5 (Additional file 2). A specificity of >50% was achieved by the SF-36 against only one tool.
6 The SF-36 did not achieve a sensitivity and specificity both >50% against any of the 162
7 HRQoL tools. For the 17 rare disease HRQoL tools, sensitivity ranged between 0-40% and
8 specificity between 5-31%. The corresponding rates for non-rare and symptom-specific tools
9 were sensitivity: 0-100% and specificity: 0-100% and sensitivity: 0-60% and specificity: 0-
10 19%, respectively.

11

12 At a prevalence of poor HRQoL of 50%, the proportion of false positives (FPR) was >50%
13 for 160/162 of the disease/condition-specific HRQoL tools and >75% for 137/162 tools
14 against the SF-36 (Figure 2). The corresponding false negative rate (FNR) was >50% for
15 160/162 tools and >75% for 159/162 tools. Decreasing the prevalence of poor HRQoL to
16 20% increased the number of studies with a FPR of >75% to 160/162 tools, while the FNR
17 was <50% for 12/162 tools. Conversely, increasing the prevalence of poor HRQoL to 80%
18 resulted in 47/162 tools having a FPR <50%, while 160/162 tools had a FNR >75%.

19

20 **Fig. 2 Misclassification (false positives and false negatives) of all disease/condition-**
21 **specific tools against the SF-36**



1

2 The figure presents misclassification of all quality of life tools (n=163) against the SF-36 by nine panels: three
 3 prevalence values (rows: 20%,50%, 80%) and three points (columns 5th percentile, median, 95th percentile).

4 Each point represents misclassification in 2 dimensions: proportions of false negatives (X-axis) and false
 5 positives (Y-axis).

6

7 *EQ-5D*

8 For the EQ-5D, a sensitivity of >50% was not reached against any of the 163

9 disease/condition-specific HRQoL tools, with the highest recorded being 50% for one tool

10 (Additional file 2). A specificity of >50% for the EQ-5D was found against 27 HRQoL tools

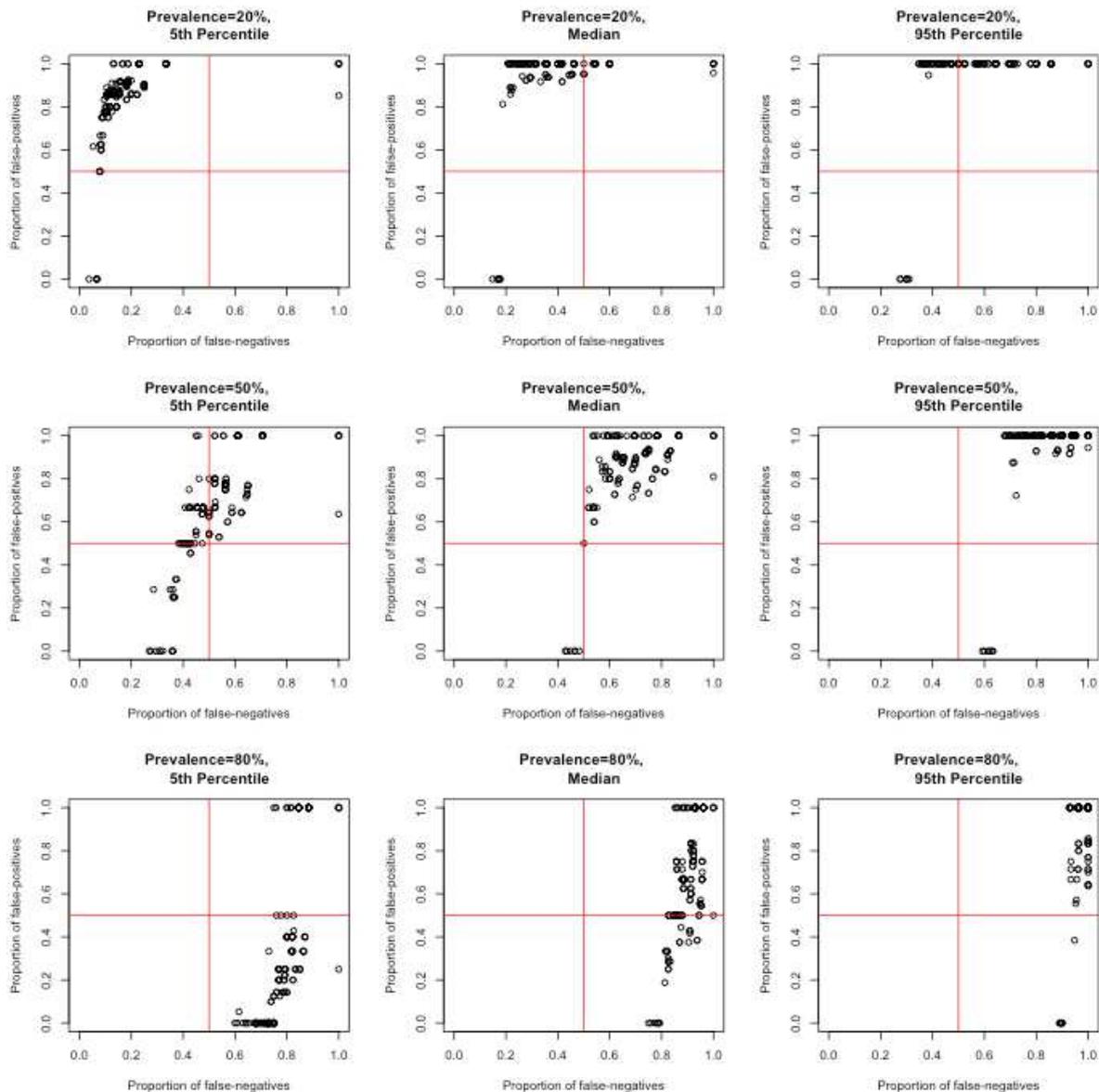
1 and >75% against 13 tools. A sensitivity and specificity both >50% was not achieved by the
2 EQ-5D against any of the tools. Sensitivity and specificity ranged from 0-22% and 29-100%,
3 respectively, against rare tools, 0-50% and 0-100% for non-rare tools, and 0-25% and 0-
4 100% for symptom-specific tools.

5

6 The FPR was >50% for 156/163 HRQoL tools and >75% for 139/163 HRQoL tools against
7 the EQ-5D, when the prevalence of poor HRQoL was set at 50% (Figure 3). Using the same
8 prevalence of poor HRQoL, 156/163 tools had a FNR of >50% and 49/163 tools a FNR of
9 >75%. A prevalence of poor HRQoL of 20% increased the number of studies with a FPR of
10 >75% to 157/163 tools, while the FNR was <50% for 138/163 tools. Increasing the
11 prevalence of poor HRQoL to 80% resulted in 48/163 tools having a FPR <50% and 162/163
12 tools having a FNR >75%.

13

14 **Fig. 3 Misclassification (false positives and false negatives) of all disease/condition-**
15 **specific tools against the EQ-5D**



1

2 The figure presents misclassification of all tools (n=163) against EQ-5D by nine panels: three prevalence values

3 (rows: 20%, 50%, 80%) and three points (columns 5th percentile, median, 95th percentile). Each point

4 represents misclassification in 2 dimensions: proportions of false negatives (X-axis) and false positives (Y-axis).

5

6 **Concordance of quality of life scores**

7 Of the disease/condition-specific HRQoL tools identified, 141 had comparative data directly

8 against a generic tool (Figure 1 & Additional file 1). For rare diseases, 11 studies (including

9 eight clinical trials) covering 10 tools provided data against the SF-36 (no comparative data

10 against the EQ-5D was identified). One-hundred and twelve studies (including 74 clinical

1 trials) covering 73 tools for non-rare diseases provided comparative data against the SF-36
 2 (84 studies) and EQ-5D (65 studies), with some publications including data for both generic
 3 tools. Seventy-one studies (43 clinical trials) of 58 symptom-specific tools were available
 4 with data against the SF-36 (66 studies) and EQ-5D (31 studies), again with some studies
 5 covering both generic tools. Overall, 63% (125/198) of the publications reported results from
 6 clinical trials.

7

8 *Strong* concordance of the SF-36 results was reported for only 27% (3/11) of studies of rare
 9 disease tools, 29% (24/84) for non-rare disease tools, and 18% (12/66) for symptom-specific
 10 tools (Table 1). Results were similar for the EQ-5D against non-rare tools, with 26% (17/65)
 11 showing a *Strong* level of concordance, although this generic tool appeared to perform
 12 marginally better than the SF-36 against symptom-specific tools (29%; 9/31). A total absence
 13 of concordance was noted for 30% (25/84) and 35% (23/65) of results from studies of non-
 14 rare disease-specific tools and 36% (24/66) and 16% (5/31) for symptom-specific tools *versus*
 15 the SF-36 and EQ-5D, respectively.

16

17 **Table 1 Concordance between health-related quality of life results for rare and non-rare**
 18 **disease and symptom-specific tools *versus* the SF-36 and EQ-5D**

Level of Concordance	Rare Disease		Non-rare Disease		Symptom-specific	
	SF-36 (n=11)	EQ-5D (n=0)	SF-36 (n=84)	EQ-5D (n=65)	SF-36 (n=66)	EQ-5D (n=31)
<i>None</i>	2 (18%)	-	25 (30%)	23 (35%)	24 (36%)	5 (16%)
<i>Moderate</i>	6 (55%)	-	35 (42%)	25 (38%)	30 (45%)	17 (55%)
<i>Strong</i>	3 (27%)	-	24 (29%)	17 (26%)	12 (18%)	9 (29%)

1 Values represent number of studies (% of total). Rare disease: affecting <1 in 2,000 population; Non-rare
 2 diseases: ≥1 in 2,000 population. Concordance: *None (No)* = significant impact on quality of life with
 3 disease/condition specific tool, but no change or the opposite impact with generic tool (or *vice versa*); *Moderate*
 4 = quality of life impact was scored in the same direction with both tools, but statistically significant with only
 5 one of them; *Strong* = results fully aligned (significant/non-significant impact in same direction)

6
 7 For studies reporting HRQoL at multiple time points, a recorded change in HRQoL over time
 8 did not result in a reclassification of concordance for rare (0/2) or non-rare disease tools
 9 (0/34) *versus* the SF-36, but did for 19% (7/36) of symptom-specific tools (Table 2). For the
 10 EQ-5D, concordance was reclassified for 9% (2/23) of studies of non-rare disease tools, but
 11 not for symptom-specific tools (0/7).

12
 13 **Table 2 Change in health-related quality of life over time and impact on concordance of**
 14 **results between rare and non-rare and symptom-specific tools *versus* the SF-36 and EQ-**
 15 **5D**

	Rare Disease		Non-rare Disease		Symptom-specific	
	SF-36 (n=2)	EQ-5D (n=0)	SF-36 (n=38)	EQ-5D (n=27)	SF-36 (n=42)	EQ-5D (n=7)
QoL change	2 (100%)	-	34 (89%)	23 (85%)	36 (86%)	7 (100%)
<i>Concordance change</i>	0/2 (0%)	-	0/34 (0%)	2/23 (9%)	7/36 (19%)	0/7 (0%)

16 n = studies with >1 time point. Rare disease: affecting <1 in 2,000 population; Non-rare diseases: ≥1 in 2,000
 17 population.

1 **Discussion**

2 An accurate measure of HRQoL is of fundamental importance when considering the clinical-
3 and cost-effectiveness of a therapy or intervention during economic evaluations/HTAs to
4 determine use within a healthcare system. Overestimating the impact on HRQoL could result
5 in excessive healthcare expenditure for minimal health gain (money which could be better
6 spent elsewhere), whereas underestimating the impact could cause unnecessary restrictions
7 on use to the detriment of patients. Selection of the appropriate tool or tools to assess HRQoL
8 is, therefore, essential. This study has found that by applying Neutral Theory generic HRQoL
9 tools, such as the SF-36 or EQ-5D, which are commonly used in HTAs, appear poorly
10 aligned with disease- or condition-specific tools.

11

12 Neither the SF-36 nor the EQ-5D achieved a sensitivity and specificity for included items
13 both >50% against any of the 162/163 disease- or condition-specific tools included in this
14 study. Even when using a high prevalence of poor HRQoL set at 80% (*i.e.* 4/5 patients with
15 this disease/condition have a notably impacted HRQoL), less than one-third of tools had a
16 FPR of <50% against the generic tools (SF-36: 29% of tools; EQ-5D: 29%). The situation
17 was worse for rare disease tools, where sensitivity ranged from 0-40% for the SF-36 and 0-
18 22% for the EQ-5D. Predicated on these results, it is unsurprising, therefore, that there were
19 low levels of concordance between HRQoL scores from the generic *versus* the
20 disease/condition-specific tools (no concordance *vs* SF-36: 18-36% of studies; *vs* EQ-5D: 16-
21 35%).

22

23 Salient limitations of this study included the necessity of having to assume a true negative
24 fraction of 0.9, as it was not possible to know if the disease/condition-specific tool fully
25 captures all relevant items or domains (*i.e.*, is completely Neutral). The measure of

1 concordance between results of the generic and disease/condition-specific tools was also
2 necessarily crude to allow for cross comparison between multiple studies of numerous
3 diseases/conditions. Importantly, however, a high number of studies (up to approximately
4 one-third) reported zero concordance between generic and disease/condition tools. It is also
5 worthy of note that, surprisingly, a full description of 29% (65/228) of the identified tools
6 could not be obtained, despite their publication in indexed journals; this is an unacceptable
7 high rate, where such descriptions should be a standard component of publication.

8

9 HRQoL tools generate scores on the basis of individual item measures – a construct. The
10 concept of 'True' HRQoL at any given time is, therefore, important. The tools generate a
11 value of observed HRQoL on a subject based on relevant items and lack of irrelevant items.
12 The principle underpinning the development and use of disease- and condition-specific tools
13 is that they are inherently more accurate than generic tools at measuring HRQoL for patients
14 with that particular disease or condition, and, thereby, closer to Neutrality. However, tools
15 have been developed for the same disease/condition that do not include all the same items
16 and domains [12, 13], which raises the question of what is the correct construct to ensure an
17 accurate assessment of HRQoL for that disease/condition. Several approaches have been
18 taken to improve the accurate assessment of HRQoL, including: the parallel use of generic
19 and disease/condition tools [14, 15]; using mapping algorithms from disease/condition-
20 specific tools to generic tools [9]; tailoring standard items to specific diseases/conditions
21 [16]; and use of bolt-on items to generic questionnaires [9]. Despite these approaches, the
22 pertinent question remains – what is acceptable accuracy for a HRQoL tool? Given the
23 importance of having an accurate measure of the impact of a therapy or intervention on
24 HRQoL, may be it is time for there to be rethink on how HRQoL is assessed and measured
25 and to consider alternative approaches to the current widespread use of generic tools.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

Conclusions

A new theory recently proposed is that of a ‘Neutral Observer’, which provides a set of principles on which a determination of the accuracy of an observation of a given construct can be based [1]. It is theorised that the “true” value of a construct can be measured by an abstract or Neutral observer who has access to a complete list of indicators that are all relevant to the empirical measurement of a construct. This Neutral Observation thereby serves as the reference against which observations using the construct can be assessed for accuracy [1]. Adoption of such an approach in the development and assessment of HRQoL tools could improve their relevance, accuracy, and utility in economic evaluations of health interventions.

1 **Supplementary Information**

2 **Additional file 1:** Details of all included and excluded tools.

3 **Additional file 2:** Results for inclusion of appropriate domains and items for all assessed tools

4 vs SF-36 and EQ-5D

5

6 **Abbreviations**

7 CI: confidence interval; EQ-5D: EuroQol five-dimensional; FNR: false negative rate; FPR:

8 false positive rate; HRQoL: Health-related quality of life; HTA: Health Technology

9 Appraisal; NICE: National Institute for Health and Care Excellence; SF-36: Medical

10 Outcomes Study Short Form family of measures; UCLA-PCI: University of California Los

11 Angeles-Prostate Cancer Index; WHO: World Health Organization;

1 Reference

- ¹ Jandhyala R. The Theory of Neutrality: a formal description and results from its testing via expert consensus in the context of clinical studies using the Jandhyala method. *BMC Med Res Methodol*. 2020 (under review).
- ² World Health Organization (WHO). WHOQOL: Measuring Quality of Life. Available at: <https://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/>. Accessed 18 December 2020.
- ³ Hays RD, Reeve BB. Measurement and Modeling of Health-Related Quality of Life. In: Heggenhougen HK, Quah S, editors. *International Encyclopedia of Public Health*. London: Academic Press; 2008. p. 241–252.
- ⁴ Ware J.E. Jr, Sherbourne C.D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care*, 30(6), 473–483.
- ⁵ RAND Corporation. 36-Item Short Form Survey (SF-36). https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html. Accessed 12 November 2019.
- ⁶ Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Annals of Medicine*. 2001;33(5):337–343.
- ⁷ EuroQol. Eq-5D. <https://euroqol.org/>. Accessed 12 November 2019.
- ⁸ National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. <https://www.nice.org.uk/process/pmg9/chapter/foreword>. Accessed 12 November 2019.
- ⁹ Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technology Assessment*. 2014;18(9):1–224.
- ¹⁰ Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile app for systematic reviews. *Syst Rev*. 2016;5(1):210.
- ¹¹ Baldovino S, Moliner AM, Taruscio D, et al. Rare Diseases in Europe: from a Wide to a Local Perspective. *Isr Med Assoc J*. 2016;18(6):359–63.
- ¹² Chen X-L, Zhong L-H, Wen Y, et al. Inflammatory bowel disease-specific health-related quality of life instruments: a systematic review of measurement properties. *Health Qual Life Outcomes*. 2017;15(1):177.
- ¹³ Marinus J, Ramaker C, van Hilten JJ, et al. Health related quality of life in Parkinson's disease: a systematic review of disease specific instruments. *J Neurol Neurosurg Psychiatry*. 2002;72(2):241–8.

-
- ¹⁴ Assari S, Lankarani MM, Montazeri A, et al. Are generic and disease-specific health related quality of life correlated? The case of chronic lung disease due to sulfur mustard. *J Res Med Sci.* 2009;14(5):285–90.
- ¹⁵ Seow LSE, Tan THG, Abdin E, et al. Comparing disease-specific and generic quality of life measures in patients with schizophrenia. *Psychiatry Res.* 2019;273:387–393.
- ¹⁶ Ware JE Jr, Gandek B, Guyer R, et al. Standardizing disease-specific quality of life measures across multiple chronic conditions: development and initial evaluation of the QOL Disease Impact Scale (QDIS[®]). *Health Qual Life Outcomes.* 2016;14:84.

Figures

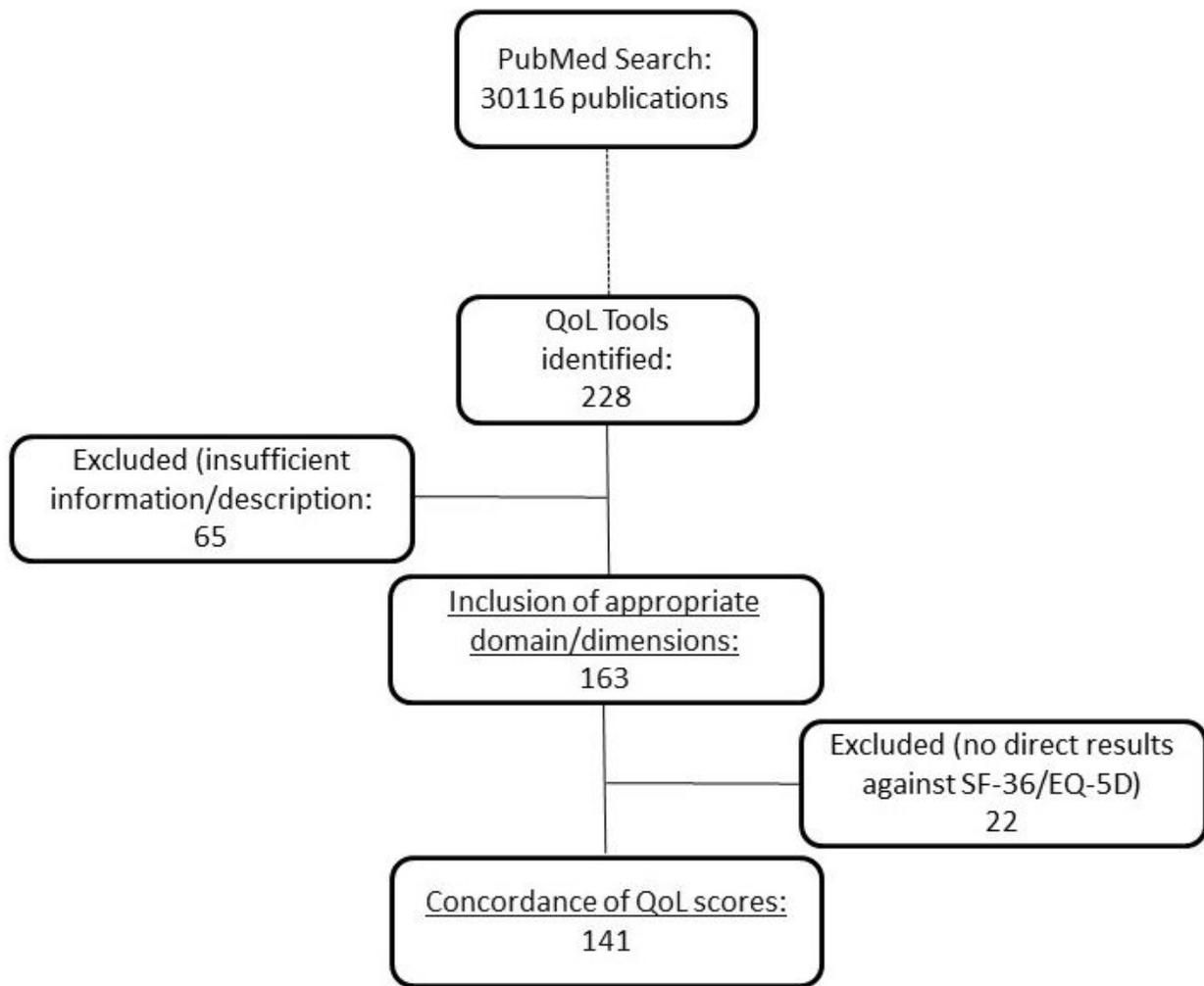


Figure 1

Overview of identification and selection of disease- and condition-specific health-related quality of life tools

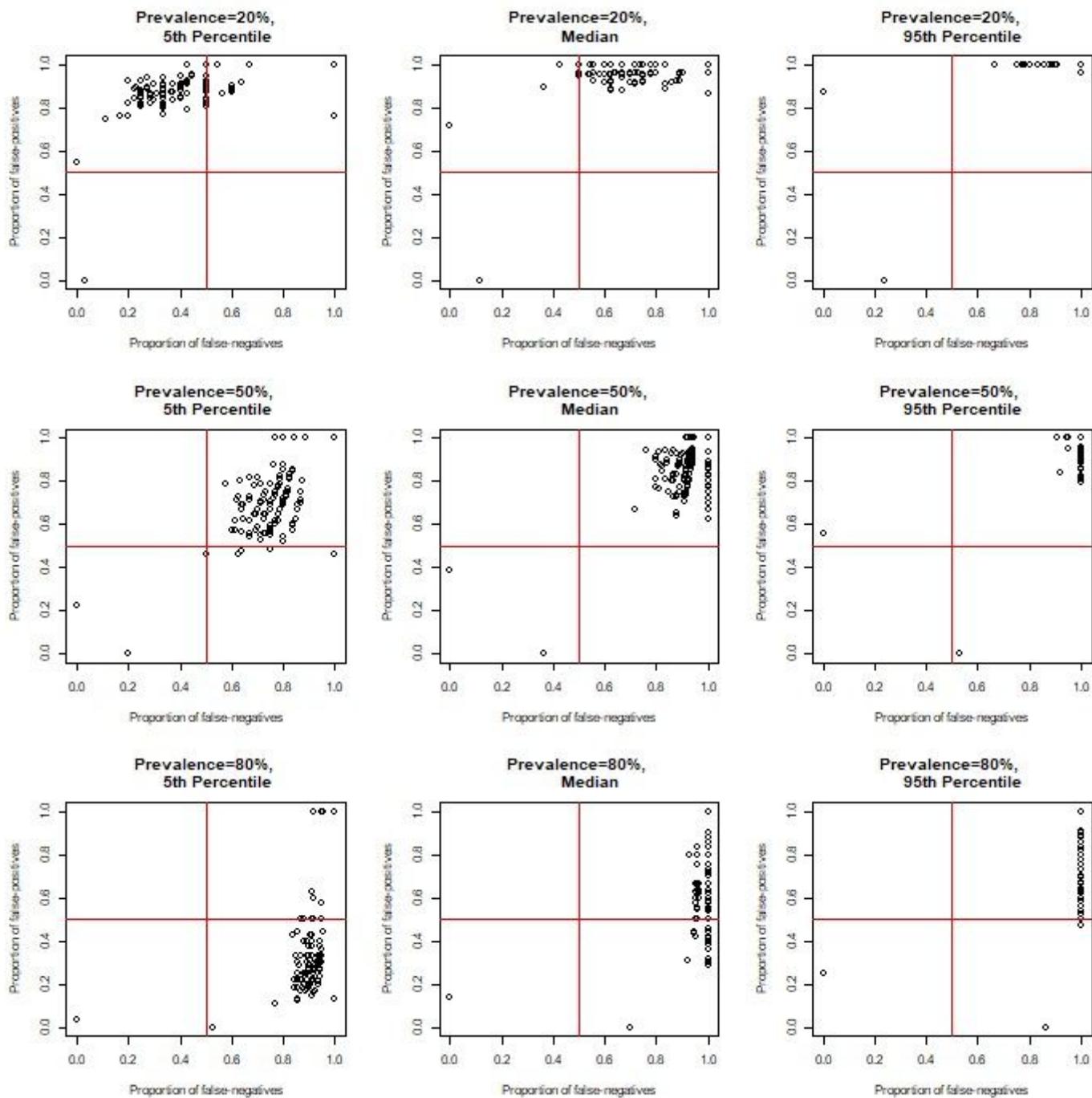


Figure 2

Misclassification (false positives and false negatives) of all disease/condition-specific tools against the SF-36. The figure presents misclassification of all quality of life tools (n=163) against the SF-36 by nine panels: three prevalence values (rows: 20%, 50%, 80%) and three points (columns: 5th percentile, median, 95th percentile). Each point represents misclassification in 2 dimensions: proportions of false negatives (X-axis) and false positives (Y-axis).

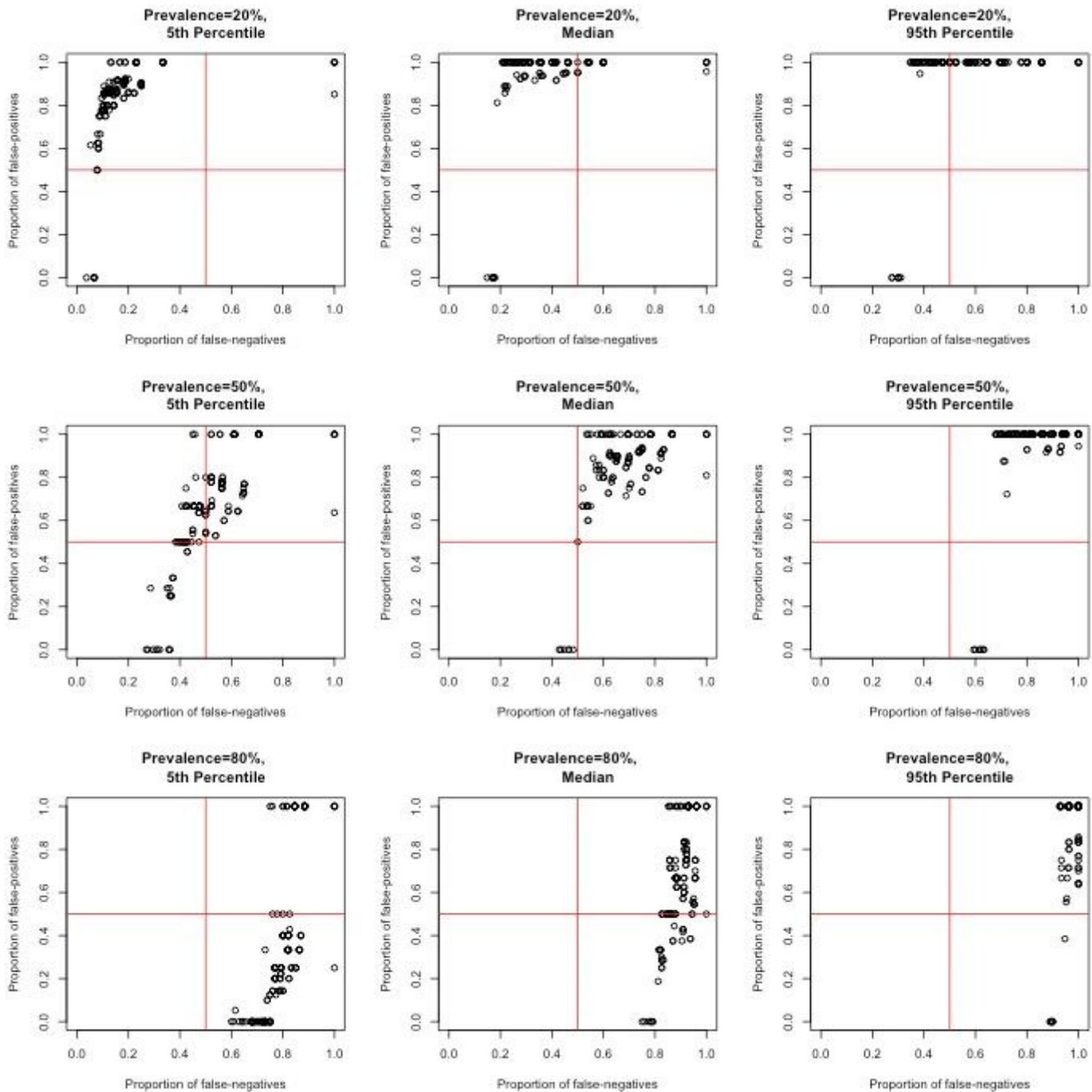


Figure 3

Misclassification (false positives and false negatives) of all disease/condition-specific tools against the EQ-5D. The figure presents misclassification of all tools (n=163) against EQ-5D by nine panels: three prevalence values (rows: 20%, 50%, 80%) and three points (columns: 5th percentile, median, 95th percentile). Each point represents misclassification in 2 dimensions: proportions of false negatives (X-axis) and false positives (Y-axis).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile1211220.xlsx](#)
- [AdditionalFile2211220.pdf](#)