**Living with Chronic Illness scale: validation in a Type 2 Diabetes Mellitus population (LW-T2DM)**

Additional material

[ADDITIONAL FILE 1. LIVING WITH T2DM SCALE 3](#_Toc43196129)

[SPANISH VERSION 3](#_Toc43196130)

[ENGLISH VERSION 4](#_Toc43196131)

[ADDITIONAL FILE 2. COSMIN ASSESSMENT 5](#_Toc43196132)

[Step 1. Evaluated measurement properties in the article 5](#_Toc43196133)

[Step 2. Determining if the statistical method used in the article are based on CTT or IRT 5](#_Toc43196134)

[Step 3. Determining if a study meets the standards for good methodological quality 6](#_Toc43196135)

[Box A. Internal consistency 6](#_Toc43196136)

[Box B. Reliability: relative measures 7](#_Toc43196137)

[Box C. Measurement error: absolute measures 8](#_Toc43196138)

[Box D. Content validity 9](#_Toc43196139)

[Box E. Structural validity 9](#_Toc43196140)

[Step 4: Determining the Generalisability of the results 10](#_Toc43196141)

[COSMIN box 1. Standards for evaluating the quality of PROM development 11](#_Toc43196142)

[Box 1. PROM development 11](#_Toc43196143)

[COSMIN box 2. Standards for evaluating the quality of content validity studies of PROMs 14](#_Toc43196144)

[Box 2. Content validity 14](#_Toc43196145)

# ADDITIONAL FILE 1. LIVING WITH T2DM SCALE

## SPANISH VERSION

|  |  |  |
| --- | --- | --- |
|  | **0 = Nunca/ nada 1 = Raramente/ casi nada 2 = A veces/ algo**  **3 = Casi siempre/ bastante 4 = Siempre/ mucho** | |
| **ACEPTACIÓN** | 1. Me disgusto cada vez que pienso que tengo **diabetes** | **0 1 2 3 4** |
| 2. Me avergüenzo de la **diabetes** y la oculto ante los demás | **0 1 2 3 4** |
| 3. Me enfada tener los síntomas de la enfermedad | **0 1 2 3 4** |
| 4. Me molesta tener que cambiar mi forma de vivir y sacrificar cosas importantes de mi vida por la enfermedad | **0 1 2 3 4** |
| **AFRONTAMIENTO** | 5. Intento hacer frente y luchar contra la enfermedad | **0 1 2 3 4** |
| 6. Tengo interés en buscar cosas que me motiven para no centrar mi vida sólo en la enfermedad *(como por ejemplo, ir al cine, compartir experiencias con otros pacientes, pasear, etc.)* | **0 1 2 3 4** |
| 7. Intento ver el lado positivo de la enfermedad | **0 1 2 3 4** |
| 8. Intento apoyarme en las cosas que son importantes en mi vida | **0 1 2 3 4** |
| 9. Tengo esperanza de que la situación mejore | **0 1 2 3 4** |
| 10. Tengo a alguien que me escucha y comprende lo que estoy viviendo con la **diabetes** | **0 1 2 3 4** |
| 11. Expreso mis sentimientos para que otras personas sepan cómo me siento y me ayuden | **0 1 2 3 4** |
| **AUTOMANEJO** | 12. Utilizo recursos para controlar y mejorar mi situación con **diabetes** | **0 1 2 3 4** |
| 13. Conozco la enfermedad y sé lo que tengo que hacer en todo momento para controlarla | **0 1 2 3 4** |
| 14. Hago todo lo que está en mis manos para que mi estado de salud no empeore | **0 1 2 3 4** |
| 15. Participo en la toma de decisiones sobre la enfermedad junto con mi médico/enfermero | **0 1 2 3 4** |
| **INTEGRACIÓN** | 16. He integrado en mi día a día la**diabetes**y todas las cosas relacionadas con ella | **0 1 2 3 4** |
| 17. A pesar de la **diabetes**, llevo una vida lo más normal posible | **0 1 2 3 4** |
| 18. Tengo en cuenta mi enfermedad a la hora de hacer actividades, tareas y/o planes | **0 1 2 3 4** |
| 19. Estoy tan interesado y tengo la misma ilusión en las actividades de ocio como antes de tener la enfermedad | **0 1 2 3 4** |
| 20. Aunque tengo **diabetes**, me siento satisfecho con mi vida | **0 1 2 3 4** |
| **ADAPTACIÓN** | 21. Gracias a la **diabetes**, ahora doy importancia a las cosas que verdaderamente tienen valor | **0 1 2 3 4** |
| 22. La **diabetes**ha cambiado mi forma de ser para bien | **0 1 2 3 4** |
| 23. He encontrado el bienestar y armonía en mi vida | **0 1 2 3 4** |
| 24. A pesar de la **diabetes**, me siento una persona igual de valiosa que otras sin la enfermedad | **0 1 2 3 4** |
| 25. A pesar de los problemas que la enfermedad genera, he encontrado un nuevo sentido en mi vida | **0 1 2 3 4** |
| 26.La **diabetes**me ha ayudado a descubrir aspectos de mí que desconocía | **0 1 2 3 4** |

**©** Leire Ambrosio. La EC-PC es una medida gratis y de acceso libre para la práctica clínica y la investigación. Antes de usarla y/o modificarla se debe pedir permiso a la autora. Para mayor detalle contacte con [lambrosio@unav.es.](mailto:lambrosio@unav.es)

## ENGLISH VERSION

|  |  |  |
| --- | --- | --- |
|  | **0 = Never/ Not at all 1 = Rarely/ A Little bit 2 = Sometimes/ somewhat**  **3 = Almost/ Always/ Quite a bit 4 = Always/ very much** | |
| **ACCEPTANCE** | 1. I get upset every time I think I have **diabetes** | **0 1 2 3 4** |
| 2. I am embarrassed about having **diabetes**and I hide it from others | **0 1 2 3 4** |
| 3. I get angry when having the illness symptoms | **0 1 2 3 4** |
| 4. It bothers me to change the way I live and sacrifice important things in my life because of the illness | **0 1 2 3 4** |
| **COPING** | 5. I try to cope and fight the disease | **0 1 2 3 4** |
| 6. I am interested in looking for things that motivate me not to focus only in **diabetes** | **0 1 2 3 4** |
| 7. I try to see the positive side of the illness | **0 1 2 3 4** |
| 8. I try to rely on things that are important in my life | **0 1 2 3 4** |
| 9. I hope the situation with **diabetes**improves | **0 1 2 3 4** |
| 10. I have someone who listens to me and understands what I am living with **diabetes** | **0 1 2 3 4** |
| 11. I express my feelings so that others know how I am feeling and can help me | **0 1 2 3 4** |
| **SELF-MANAGEMENT** | 12. I make use of resources to control and improve my situation with **diabetes** | **0 1 2 3 4** |
| 13. I know this illness and I always know what I have to do to control it | **0 1 2 3 4** |
| 14. I do everything possible to prevent my health from getting worse | **0 1 2 3 4** |
| 15. I am involved in decision-making about my illness with my doctor/nurse | **0 1 2 3 4** |
| **INTEGRATION** | 16. I have integrated **diabetes**in my daily living and all things related to it | **0 1 2 3 4** |
| 17. Despite **diabetes**, I lead a life as normal as possible | **0 1 2 3 4** |
| 18. I take my illness into account when it comes to doing activities, tasks, and/or making plans | **0 1 2 3 4** |
| 19. I am as interested and I have the same illusion in leisure activities as before having **diabetes** | **0 1 2 3 4** |
| 20. Although I have **diabetes**I feel satisfied with my life | **0 1 2 3 4** |
| **ADJUSTMENT** | 21. Now I value the things that really matter thanks to **diabetes** | **0 1 2 3 4** |
| 22. **Diabetes**has changed the way I am for the better | **0 1 2 3 4** |
| 23. I have found well-being and harmony in my life | **0 1 2 3 4** |
| 24. Despite having **diabetes**, I feel that I am just as valuable a person as others who don’t have the illness | **0 1 2 3 4** |
| 25. Despite the problems the illness creates, I have found new meaning in my life | **0 1 2 3 4** |
| 26.**Diabetes**has helped me discover things about myself that I didn’t know before | **0 1 2 3 4** |

**©** Leire Ambrosio. The LW-CI Scale is a free and open Access instrument to be used in clinical practice and research. However, before using or modifying it, main author should be contacted in the following email address: [lambrosio@unav.es.](mailto:lambrosio@unav.es)

# ADDITIONAL FILE 2. COSMIN ASSESSMENT

## Step 1. Evaluated measurement properties in the article

|  |  |  |
| --- | --- | --- |
| X | Internal consistency | Box A |
| X | Reliability | Box B |
| X | Measurement error | Box C |
| X | Content validity | Box D |
| X | Structural validity | Box E |
|  | Hypotheses testing | Box F |
|  | Cross-cultural validity | Box G |
|  | Criterion validity | Box H |
|  | Responsiveness | Box I |
|  | Interpretability | Box J |

## Step 2. Determining if the statistical method used in the article are based on CTT or IRT

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Box General requirements for studies that applied Item Response Theory (IRT) models** | | | | | | |
|  |  | **yes** |  | **no** |  | **?** |  |
| 1 | Was the IRT model used adequately described? e.g. One Parameter Logistic Model (OPLM), Partial Credit Model (PCM), Graded Response Model (GRM) |  |  | X |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| 2 | Was the computer software package used adequately described? e.g. RUMM2020, WINSTEPS, OPLM, MULTILOG, PARSCALE, BILOG, NLMIXED |  |  | X |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| 3 | Was the method of estimation used adequately described? e.g. conditional maximum likelihood (CML), marginal maximum likelihood (MML) |  |  | X |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| 4 | Were the assumptions for estimating parameters of the IRT model checked? e.g. unidimensionality, local independence, and item fit (e.g. differential item functioning (DIF)  \*The model used in the LW-T2DM scale is based on Classical Test Theory |  |  |  |  | X |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

## Step 3. Determining if a study meets the standards for good methodological quality

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Box A. Internal consistency |  |  |  |  | |  | | |  | | |  | | |  | |
|  |  | **yes** |  | **no** |  | | **?** | | |  | | |  | | |  | |
| 1 | Does the scale consist of effect indicators, i.e. is it based on a reflective model? | X |  |  |  | |  | | |  | | |  | | |  | |
| *Design requirements* | | **yes** |  | **no** |  | | **?** | | |  | | |  | | |  | |
| 2 | Was the percentage of missing items given? | X |  |  |  | |  | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
| 3 | Was there a description of how missing items were handled? | X |  |  |  | |  | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
| 4 | Was the sample size included in the internal consistency analysis adequate? | X |  |  |  | |  | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
| 5 | Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT |  |  | X |  | |  | | |  | | |  | | |  | |
|  | model applied? |  |  |  |  | |  | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
| 6 | Was the sample size included in the unidimensionality analysis adequate? |  |  | X |  | |  | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
| 7 | Was an internal consistency statistic calculated for each (unidimensional) (sub)scale | X |  |  |  | |  | | |  | | |  | | |  | |
|  | separately? |  |  |  |  | |  | | |  | | |  | | |  | |
| 8 | Were there any important flaws in the design or methods of the study? |  |  | X |  | |  | | |  | | |  | | |  | |
| *Statistical methods* | | **yes** |  | **no** |  | | **NA** | | |  | | |  | | |  | |
| 9 | for Classical Test Theory (CTT): Was Cronbach’s alpha calculated? | X |  |  |  | |  | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
| 10 | for dichotomous scores: Was Cronbach’s alpha or KR-20 calculated? |  |  |  |  | | X | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
| 11 | for IRT: Was a goodness of fit statistic at a global level calculated? e.g. χ2, reliability |  |  |  |  | | X | | |  | | |  | | |  | |
|  | coefficient of estimated latent trait value (index of (subject or item) separation) |  |  |  |  | |  | | |  | | |  | | |  | |
|  |  |  |  |  |  | |  | | |  | | |  | | |  | |
|  | Box B. Reliability: relative measures **(including test-retest reliability, inter-rater reliability and intra-rater reliability)** | | | | | | | | | | | | | | | | |
| *Design requirements* | | | | | | **yes** | |  | **no** | |  | **?** | |  |  | |  |
| 1 | Was the percentage of missing items given? | | | | |  | |  | X | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 2 | Was there a description of how missing items were handled? | | | | |  | |  | X | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 3 | Was the sample size included in the analysis adequate? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 4 | Were at least two measurements available? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 5 | Were the administrations independent? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 6 | Was the time interval stated? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 7 | Were patients stable in the interim period on the construct to be measured? | | | | |  | |  |  | |  | X | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 8 | Was the time interval appropriate? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 9 | Were the test conditions similar for both measurements? e.g. type of | | | | | X | |  |  | |  |  | |  |  | |  |
|  | administration, environment, instructions | | | | |  | |  |  | |  |  | |  |  | |  |
| 10 | Were there any important flaws in the design or methods of the study? | | | | |  | |  | X | |  |  | |  |  | |  |
| *Statistical methods* | | | | | | **yes** | |  | **no** | |  | **NA** | |  | **?** | |  |
| 11 | for continuous scores: Was an intraclass correlation coefficient (ICC) calculated? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 12 | for dichotomous/nominal/ordinal scores: Was kappa calculated? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 13 | for ordinal scores: Was a weighted kappa calculated? | | | | | X | |  |  | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |
| 14 | for ordinal scores: Was the weighting scheme described? e.g. linear, quadratic | | | | |  | |  | X | |  |  | |  |  | |  |
|  |  | | | | |  | |  |  | |  |  | |  |  | |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Box C. Measurement error: absolute measures |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *Design requirements* | | **yes** |  | **no** |  | **?** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 1 | Was the percentage of missing items given? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 2 | Was there a description of how missing items were handled? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 3 | Was the sample size included in the analysis adequate? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 4 | Were at least two measurements available? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 5 | Were the administrations independent? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 6 | Was the time interval stated? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 7 | Were patients stable in the interim period on the construct to be measured? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 8 | Was the time interval appropriate? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 9 | Were the test conditions similar for both measurements? e.g. type of | X |  |  |  |  |  |  |  |
|  | administration, environment, instructions |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 10 | Were there any important flaws in the design or methods of the study? |  |  | X |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *Statistical methods* | | **yes** |  | **no** |  | **?** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 11 | for CTT: Was the Standard Error of Measurement (SEM), Smallest Detectable Change | X |  |  |  |  |  |  |  |
|  | (SDC) or Limits of Agreement (LoA) calculated? |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Box D. Content validity **(including face validity)** |  |  |  |  |  |  |  |  |
| *General requirements* | | **yes** |  | **no** |  | **?** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 1 | Was there an assessment of whether all items refer to relevant aspects of the | X |  |  |  |  |  |  |  |
|  | construct to be measured? |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 2 | Was there an assessment of whether all items are relevant for the study | X |  |  |  |  |  |  |  |
|  | population? (e.g. age, gender, disease characteristics, country, setting) |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 3 | Was there an assessment of whether all items are relevant for the purpose of the |  |  |  |  | X |  |  |  |
|  | measurement instrument? (discriminative, evaluative, and/or predictive) |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 4 | Was there an assessment of whether all items together comprehensively reflect the | X |  |  |  |  |  |  |  |
|  | construct to be measured? |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 5 | Were there any important flaws in the design or methods of the study? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Box E. Structural validity |  |  |  |  |  |  |  |  |
|  |  | **yes** |  | **no** |  | **?** |  |  |  |
| 1 | Does the scale consist of effect indicators, i.e. is it based on a reflective model? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *Design requirements* | | **yes** |  | **no** |  | **?** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 2 | Was the percentage of missing items given? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 3 | Was there a description of how missing items were handled? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 4 | Was the sample size included in the analysis adequate? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 5 | Were there any important flaws in the design or methods of the study? |  |  | X |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *Statistical methods* | | **yes** |  | **no** |  | **NA** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 6 | for CTT: Was exploratory or confirmatory factor analysis performed? |  |  | X |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 7 | for IRT: Were IRT tests for determining the (uni-) dimensionality of the items |  |  |  |  | X |  |  |  |
|  | performed? |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

## Step 4: Determining the Generalisability of the results

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Box Generalisability** |  |  |  |  |  |  |  |  |
|  |  | **yes** |  | **no** |  | **NA** |  |  |  |
| Was the sample in which the HR-PRO instrument was evaluated adequately described? | |  |  |  |  |  |  |  |  |
| In terms of: | |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 1 | median or mean age (with standard deviation or range)? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 2 | distribution of sex? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 3 | important disease characteristics (e.g. severity, status, duration) and description |  |  |  |  | X |  |  |  |
|  | of treatment? |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 4 | setting(s) in which the study was conducted? e.g. general population, primary care | X |  |  |  |  |  |  |  |
|  | or hospital/rehabilitation care |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 5 | countries in which the study was conducted? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 6 | language in which the HR-PRO instrument was evaluated? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 7 | Was the method used to select patients adequately described? e.g. convenience, | X |  |  |  |  |  |  |  |
|  | consecutive, or random |  |  |  |  |  |  |  |  |
|  |  | **yes** |  | **no** |  | **?** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 8 | Was the percentage of missing responses (response rate) acceptable? | X |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

## COSMIN box 1. Standards for evaluating the quality of PROM development

Part 1a (PROM design) and part 1b (Cognitive interview study or other pilot test) both need to be completed for each PROM because all standards of part 1a and part 1b will be

included in the final rating of the quality of the PROM development. However, if a cognitive interview study or other pilot test was not performed, only the first standard in part

1b needs to be completed and the rest of the box can be skipped. For rating the standards, the “worst score counts” method is used. A total rating for the box can be obtained

by taking the lowest rating of any standard in the box. It is also possible to obtain total ratings for different parts of the boxes by taking the lowest rating of any standard of that

part of the box.

### Box 1. PROM development

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1a. PROM design** | |  |  |  |  |  |  |  |  |  |
| *General design requirements* | | **Very Good** |  | **Adequate** |  | **Doubful** |  | **Inadequate** |  | **Not aplicable** |
| 1 | Is a clear description provided of the construct to be measured? | Construct clearly described |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 2 | Is the origin of the construct clear: was a theory, conceptual framework or disease model used or clear rationale provided to define the construct to be measured? | Origin of the construct clear |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 3 | Is a clear description provided of the target population for which the PROM was developed? | Target population clearly described |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 4 | Is a clear description provided of the context of use | Context of use clearly described |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *General design requirements* | | **Very Good** |  | **Adequate** |  | **Doubful** |  | **Inadequate** |  | **Not aplicable** |
| 5 | Was the PROM development study performed in a sample representing the target population for which the PROM was developed? | Study performed in a sample representing the target population |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *Concept elicitation (relevance and comprehensiveness)* | |  |  |  |  |  |  |  |  |  |
| 6 | Was an appropriate qualitative data collection method used to identify relevant items for a new PROM? |  |  | Assumable that the qualitative method was appropriate and suitable for the construct and study population, but not clearly described |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 7 | Were skilled group moderators/interviewers used? |  |  |  |  |  |  |  |  | Not applicable |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 8 | Were the group meetings or interviews based on an appropriate topic or interview guide? |  |  |  |  |  |  |  |  | Not applicable |
|  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *General design requirements* | | **Very Good** |  | **Adequate** |  | **Doubful** |  | **Inadequate** |  | **Not aplicable** |
| 9 | Were the group meetings or interviews recorded and transcribed verbatim? |  |  |  |  |  |  |  |  | Not applicable |
|  |  |  |  |  |  |  |  |  |  |  |
| 10 | Was an appropriate approach used to analyse the data? |  |  |  |  |  |  |  |  | Not applicable |
|  |  |  |  |  |  |  |  |  |  |
| 11 | Was at least part of the data coded independently? |  |  |  |  |  |  |  |  | Not applicable |
|  |  |  |  |  |  |  |  |  |  |  |
| 12 | Was data collection continued until saturation was reached? |  |  |  |  |  |  |  |  | Not applicable |
|  |  |  |  |  |  |  |  |  |  |  |
| 13 | For quantitative studies (surveys): was the sample size appropriate? | ≥100 |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1b. Cognitive interview study or other pilot test** | |  |  |  |  |  |  |  |  |  |
|  |  | **Very Good** |  | **Adequate** |  | **Doubful** |  | **Inadequate** |  | **Not aplicable** |
| 14 | Was a cognitive interview study or other pilot test conducted? | YES |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *General design requirements* | |  |  |  |  |  |  |  |  |  |
| 15 | Was the cognitive interview study or other pilot test performed in a sample representing the target population? | Study performed in a sample representing the target population |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Comprehensibility* | |  |  |  |  |  |  |  |  |  |
| 16 | Were patients asked about the comprehensibility of the PROM? |  |  |  |  | Not clear **(SKIP standards 17-35)** |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

## COSMIN box 2. Standards for evaluating the quality of content validity studies of PROMs

### Box 2. Content validity

Only those parts of the box need to be completed for which information is available. For example, if a content validity study was not performed in professionals, parts 2d and 2e do

not need to be completed. If patients were included in a content validity study, but they were only asked about comprehensibility of the PROM items, sections 2a and 2b do not need

to be completed.  
For rating the standards, the “worst score counts” method is used. A total score for the box can be obtained by taking the lowest rating of any standard in the box. It is also possible to

obtain total ratings for different parts of the boxes by taking the lowest rating of any standard of that part of the box.

Content validity study was not performed in professionals or patients.