



بِسْمِ اللَّهِ

الرَّحْمَنِ

الرَّحِيمِ

معرفی فرم نقد ابزار "CosMIN"

شاخص های انتخاب ابزار مناسب برای انجام پژوهش در علوم سلامت:

معرفی فهرست COSMIN



Enhancing the QUALity and Transparency Of health Research

EQUATOR resources in [Portuguese](#) | [Spanish](#)

- [Home](#)
- [Library](#)
- [Toolkits](#)
- [Courses & events](#)
- [News](#)
- [Blog](#)
- [Librarian Network](#)
- [About us](#)
- [Contact](#)

Your one-stop-shop for writing and publishing high-impact health research
 find reporting guidelines | improve your writing | join our courses | run your own training course | enhance your peer review | implement guidelines

Library for health research reporting

The Library contains a comprehensive searchable database of reporting guidelines and also links to other resources relevant to research reporting.

- [Search for reporting guidelines](#)
- [Not sure which reporting guideline to use?](#)
- [Reporting guidelines under development](#)
- [Visit the library for more resources](#)

Reporting guidelines for main study types

Randomised trials	CONSORT	Extensions	Other
Observational studies	STROBE	Extensions	Other
Systematic reviews	PRISMA	Extensions	Other
Case reports	CARE	Extensions	Other
Qualitative research	SRQR	COREQ	Other
Diagnostic / prognostic studies	STARD	TRIPOD	Other
Quality improvement studies	SQUIRE		Other
Economic evaluations	CHEERS		Other
Animal pre-clinical studies	ARRIVE		Other
Study protocols	SPIRIT	PRISMA-P	Other
Clinical practice guidelines	AGREE	RIGHT	Other

EQUATOR Oncology



Resources supporting complete, accurate and transparent research



12/22/2021

Search for reporting guidelines

Use your browser's Back button to return to your search results



COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures

Reporting guideline provided for?
(i.e. exactly what the authors state in the paper)

Reporting of primary studies of the measurement properties of patient-reported outcome measures (PROMs).

Full bibliographic reference

Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. Qual Life Res. 2021.

Language

English

PubMed ID

[33818733](#)



Reporting guidelines for main study types

Randomised trials	CONSORT	Extensions
Observational studies	STROBE	Extensions
Systematic reviews	PRISMA	Extensions
Study protocols	SPIRIT	PRISMA-P
Diagnostic/prognostic studies	STARD	TRIPOD
Case reports	CARE	Extensions
Clinical practice guidelines	AGREE	RIGHT
Qualitative research	SRQR	COREQ
Animal pre-clinical studies	ARRIVE	
Quality improvement studies	SQUIRE	Extensions

An introduction to COSMIN

“**CO**nsensus-based **S**tandards for the selection of health **M**easurement **IN**struments”

in

Lidwine B. Mokkink

Caroline B Terwee

Donald L Patrick

Jordi Alonso

Paul W Stratford

Dirk L Knol

Lex M Bouter

Henrica CW de Vet

PROMs

استانداردهای مبتنی بر اجماع جهت انتخاب ابزارهای اندازه گیری سلامت

اهداف طراحی چک لیست “CosMIN”

► تعیین ویژگی های روانسنجی که برای هر ابزار باید مورد بررسی قرار گیرد و ارائه تعریفی از این ویژگی ها

► ارائه معیارهایی جهت بررسی مناسب بودن این ویژگی ها

اهداف و کاربردهای فرم های نقد:

➤ در طراحی مطالعات

➤ قبل از استفاده از یک ابزار

➤ ارائه گزارش

➤ بررسی نقادانه مطالعات

در تهیه ی این فهرست،

از **75** نفر از متخصصین در زمینه روانشناسی، اپیدمیولوژی، آمار و پزشکی و نیز سازمانهایی نظیر "انجمن بین المللی کیفیت زندگی

International Society Quality of Life

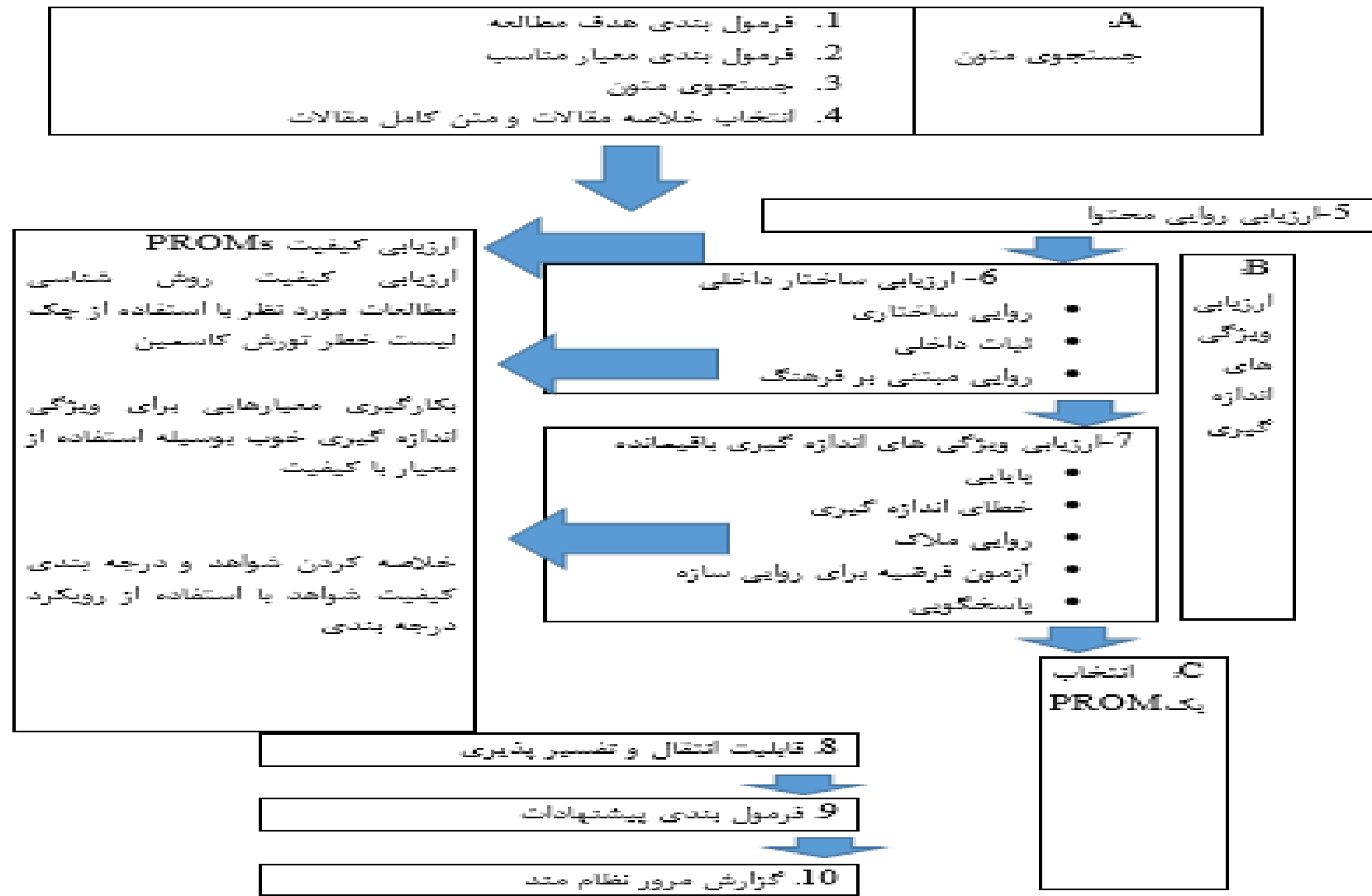
و "گروه پژوهش اروپایی برآیندهای سلامت

European Research Group on Health Outcomes

که دارای کتب و مقالاتی در خصوص ویژگیهای روان سنجی و ابزارهای اندازه گیری سلامت بودند، از نقاط مختلف جهان شامل آمریکای شمالی، اروپا، استرالیا و آسیا، جهت شرکت در بحث های گروهی دعوت به عمل آمد



Figure 1. The COSMIN taxonomy



جدول شماره 2- ده مرحله برای هدایت مرور نظام مند PROM

ساختار کلی چک لیست “CosMIN”

فهرست طراحی شده توسط این گروه، به ارزیابی کیفیت یک پژوهش، بر اساس ویژگیهای روان سنجی ابزارهای استفاده شده در آن پژوهش، می پردازد

فهرست COSMIN دارای 12 حیطه مجزا است که در 4 مرحله به بررسی ویژگیهای روان سنجی می پردازد در مرحله ی **نخست**، تشخیص این که کدام ویژگیها باید در یک مطالعه ارزیابی شوند، مطرح می گردد. این ویژگیها در ده حیطه بررسی می شوند.

مرحله ی دوم مربوط به تشخیص این مساله است که آیا روشهای آماری استفاده شده بر اساس نظریه **کلاسیک آزمون Classical Test Theory** است یا **نظریه سؤال-پاسخ Item Response Theory**

شاخصهای مربوطه، صرفا در یک حیطه بررسی می گردد نظریه های کلاسیک آزمون و سؤال- پاسخ، در طراحی و تحلیل پرسشنامه ها کاربرد زیادی دارند.



• Constructing Scale Theories

- Classic Test theory
- Item Response Theory
- Rasch Analysis
- Multidimensional Scaling
- Unfolding Theory

Classic Test theory

- توسط Nunnally در سال 1978 بیان شد.
- از این تئوری بیش از سایر تئوری‌ها استفاده می‌شود.
- دارای مبانی ریاضی و آماری نسبتاً ساده‌تری است.
- توسط نرم‌افزارهای آماری موجود قابلیت تحلیل دارد.
- ابزارهای پرستاری عمدتاً بر اساس این تئوری طراحی می‌شوند.



Classic Test theory

هدف کلی درک و بهبود روایی آزمون ها است.
هر فردی یک نمره واقعی دارد،
نمره مشاهده شده برابر است با
مجموع نمره واقعی و خطا.



Item Response Theory

- ▶ Latent Trait Theory
- ▶ تاکید بر پاسخ تک تک سوالات
- ▶ احتمال پاسخ درست دادن به یک سوال تابع توانایی پاسخ دهنده در آن صفت است.

Rasch Analysis

- ▶ در این مدلها نیز مانند مقیاس های گاتمن فرض براین است که همه سؤال ها و مواد يك آزمون که يك سازه را اندازه گیری مي کنند، يك نوع رابطه مرتب شده را تشكيل مي دهند.
- ▶ يك آزمون ممکن است دارای همسانی درونی مرتب شده اي باشد، حتي اگر مجموعه سؤال های آن همبستگی بالایی با هم نداشته باشند

Rash Analysis

► يك سؤال دشوار مي‌تواند پاسخ به سؤالهاي با دشواری کمتر را پیش‌بینی کند اما عکس

آن امکان‌پذیر نیست

► این مدل را می‌توان در مقیاسهای لیکرت، درجه‌بندی و نیز سؤالهای مربوط به اندازه‌گیریهای

ترتیبی که در آنها نمره‌های متوالی بالاتر بیانگر سطح فزاینده پیشرفت و توانمندی است به

کار برد.

Multidimensional Scaling

- با تجسم اطلاعات به دنبال کشف تشابهات وعدم تشابهات در داده ها هستند و با یک الگوریتم دارای ماتریکس عبارت- عبارت مشابه شروع می شود. می توان آنرا یک یک راه دیگر برای تحلیل عاملی در نظر گرفت

آشکارسازی *Unfolding*

- ▶ برای داده های مخالف - موافق هستند *Unfolding* مدل های
- ▶ برای سنجش نگرش

تمرکز نظریه ی کلاسیک آزمون بیشتر بر روی عملکرد آزمون و نمره ای است که در آزمون کسب شده و هدف آن جستجوی ارتباط و همبستگی بین نمره واقعی و نمرات آزمون است.

در حالی که تمرکز نظریه سؤال- پاسخ بر **عملکرد سؤال و توانایی** پاسخ دهندگان و **نحوه برخورد پاسخ دهندگان با سؤال بوده و** به بررسی این موضوع می پردازد که چگونه می توان **ارتباطی معنادار بین عملکرد افراد و توانایی** هایی که به وسیله آزمون سنجیده می شود، برقرار نمود.

مرحله ی سوم، شامل بررسی روش رعایت استاندارد در تعیین ویژگیهای روان سنجی بوده که شاخص های این مرحله در 10 حیطه مجزا بررسی می شود. در این قسمت، به توصیف و ارزیابی مواردی پرداخته می شود که در مرحله ی نخست، صرفاً به آنها اشاره شده است.

در نهایت در **مرحله ی چهارم**، الزامات عمومی برای تعمیم پذیری نتایج در یک حیطه مطرح می شود.

The checklist contains **12 boxes**.

- **Nine** of these boxes contain *standards for the included measurement properties*:
 - internal consistency (box A),
 - reliability (box B),
 - measurement error (box C),
 - content validity (including face validity)(box D),
 - construct validity (i.e. structural validity (box E),
 - hypotheses testing (box F),
 - cross-cultural validity (box G),
 - criterion validity (box H),
 - responsiveness (box I).
- **One** box contains *standards for studies on interpretability (box J)*.
- **two** boxes are included in the checklist that contain *general requirements: (IRT and Generalizability boxes)*

Domain	Term		Definition
	Measurement property	Aspect of a measurement property	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health related-patient reported outcomes (HR-PRO) (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true' [†] differences between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
	Face validity		The degree to which (the items of) an HR-PRO instrument indeed looks as though they are an adequate reflection of the construct to be measured

جدول شماره 1: تعاریف ابعاد کاسمین، ویژگی های اندازه گیری و جنبه های ویژگی های اندازه گیری

تعریف		واژه	
جنبه یک ویژگی اندازه گیری	ویژگی اندازه گیری	اندازه	بعد
درجه ای که در آن اندازه گیری از خطای اندازه گیری آزاد است.			پایایی
میزان نمرات برای بیمارانی که تغییر نکردند و در اندازه گیری های مکرر تحت چندین شرایط یکسان هستند: به عنوان مثال استفاده از مجموعه آیتم های متفاوت از PROM یکسان (ثبات داخلی) انجام آزمون - باز آزمون توسط افراد مختلف روی موارد یکسان (بین رتبه دهندگان) یا توسط افراد یکسان (به عنوان مثال رتبه دهندگان یا پاسخ دهندگان) روی موارد متفاوت (intra-rater)			پایایی (تعریف جامع)

دستور العمل تكميل چك ليست

“CosMIN”

INSTRUCTIONS FOR COMPLETING THE COSMIN CHECKLIST

Mark the properties that have been assessed in the article.

- | | |
|--|--------------------------|
| A. Internal consistency | <input type="checkbox"/> |
| B. Reliability | <input type="checkbox"/> |
| C. Measurement error | <input type="checkbox"/> |
| D. Content validity
(including face validity) | <input type="checkbox"/> |
| Construct validity | |
| E. Structural validity | <input type="checkbox"/> |
| F. Hypotheses-testing | <input type="checkbox"/> |
| G. Cross-cultural validity | <input type="checkbox"/> |
| H. Criterion validity | <input type="checkbox"/> |
| I. Responsiveness | <input type="checkbox"/> |
| J. Interpretability | <input type="checkbox"/> |

Step 1

Are IRT methods used in the article?

Yes

No

Complete IRT box

Step 2

Complete for each property you marked in step 1 the corresponding box A to J

Step 3

Complete for each property you marked in step 1 the Generalisability box

Step 4

جدول ۱- چهار گام تکمیل فهرست COSMIN

حیطه		ویژگی	گام	
روایی محتوا	روایی	تعیین ویژگیهای روان سنجی که در مطالعات مورد ارزیابی قرار گرفته اند.	۱	
روایی ساختار آزمون فرضیه روایی بین فرهنگی				روایی سازه
روایی ملاکی				پایایی
همسانی درونی پایایی خطای اندازه گیری	پاسخگویی			
	تفسیرپذیری			
الزامات عمومی مطالعاتی که نظریه ی پاسخ سؤال را به کار برده اند.		تکمیل جدول مربوط در این خصوص که آیا نظریه ی سؤال- پاسخ در مطالعات استفاده شده است یا خیر.	۲	
تشریح ۱۰ حیطه ی مطرح شده در گام اول		تکمیل جداول مربوط به هر یک از ویژگیهای روان سنجی که در گام اول مشخص شد.	۳	
تعمیم پذیری		تکمیل جدول تعمیم پذیری برای یک از ویژگیهای روان سنجی که در گام اول مشخص شد.	۴	

<p>روایی محتوا (شامل روایی صوری): آیا ارزیابی مبتنی بر اینکه تمام گویه ها به جنبه های مرتبطی از سازه مورد بررسی اشاره می کنند، انجام شده است؟ آیا ارزیابی اینکه تمام گویه ها مرتبط با جامعه پژوهش هستند (سن، جنس، ویژگیهای بیماری، کشور و مکان)، انجام شده است؟ آیا ارزیابی اینکه تمام گویه ها مرتبط با اهداف ابزار اندازه گیری هستند، انجام شده است؟ آیا ارزیابی اینکه تمام گویه ها در ارتباط با هم، منعکس کننده ی سازه ی مورد اندازه گیری هستند، انجام شده است؟</p>	
<p>روایی ساختار: آیا ابزار شامل شاخص های مناسب است؟ آیا در نظریه های کلاسیک آزمون، تحلیل عامل اکتشافی یا تاییدی انجام شده است؟ آیا در نظریه ی سؤال - پاسخ، آزمون مرتبط برای تشخیص بعد یا ابعاد گویه ها انجام شده است؟</p>	
<p>آزمون فرضیه: آیا فرضیه ای مبنی بر وجود ارتباط یا تفاوت میانگین ها قبل از جمع آوری داده ها شکل گرفته است؟ آیا فرضیه شامل جهت مورد انتظار در همبستگی یا تفاوت میانگین ها می باشد؟ آیا فرضیه شامل میزان مطلق یا نسبی مورد انتظار همبستگی یا تفاوت میانگین ها می باشد؟ آیا در روایی همگرا، توصیفی کامل از ابزار معیار، ارائه شده است؟</p>	<p>روایی سازه</p>
<p>روایی مبتنی بر فرهنگ: آیا تخصص هر یک از افرادی که در فرایند ترجمه شرکت داشتند، به طور کامل توصیف شده است؟ آیا مترجمین به صورت مستقل از هم کار کرده اند؟ کدام گویه ها ترجمه و ترجمه برگردان شده است؟ آیا به اندازه کافی در خصوص چگونگی رفع تفاوت بین نسخه اصلی و ترجمه شده، توضیح داده شده است؟ آیا ترجمه توسط گروهی (به عنوان مثال طراحان اولیه ابزار) مورد بازبینی قرار گرفته است؟ آیا ابزار جهت تفسیر، تطابق فرهنگی نسخه ی ترجمه شده و سهولت درک، آزمون مقدماتی انجام شده است؟ آیا نمونه ها به غیر از زبان و بستر فرهنگی در تمام ویژگیها، مشابه بودند؟ آیا عملکرد متفاوت گویه در دو گروه زبان اصلی و ترجمه شده، بررسی شده است؟</p>	<p>روایی</p>

Box A. Internal consistency

	yes	no	?
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Design requirements</i>			
	yes	no	?
2 Was the percentage of missing items given?	<input type="checkbox"/>	<input type="checkbox"/>	
3 Was there a description of how missing items were handled?	<input type="checkbox"/>	<input type="checkbox"/>	
4 Was the sample size included in the internal consistency analysis adequate?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	<input type="checkbox"/>	<input type="checkbox"/>	
6 Was the sample size included in the unidimensionality analysis adequate?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7 Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8 Were there any important flaws in the design or methods of the study?	<input type="checkbox"/>	<input type="checkbox"/>	
<i>Statistical methods</i>			
	yes	no	NA
9 for Classical Test Theory (CTT): Was Cronbach's alpha calculated?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10 for dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11 for IRT: Was a goodness of fit statistic at a global level calculated? e.g. χ^2 , reliability coefficient of estimated latent trait value (index of (subject or item) separation)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

روایی ملاکی:

آیا ملاک به کار برده شده می تواند به عنوان یک استاندارد طلایی منطقی، محسوب می گردد؟
آیا برای نمرات دو ارزشی، حساسیت و ویژگی تعیین شده است؟

همسانی درونی:

آیا ابزار شامل نشانگرهای مناسب است؟
آیا تک بعدی بودن ابزار چک شده است؟ به طور مثال آیا تحلیل عاملی انجام شده است؟
آیا همسانی درونی برای ابزار تک بعدی و نیز برای هر یک از ابعاد در (ابزار) چند بعدی، به صورت مجزا محاسبه شده است؟
آیا در نظریه های کلاسیک آزمون، آلفای کرونباخ محاسبه شده است؟
آیا در عبارات سؤالات دوازدهی، آلفای کرونباخ یا کوادر ریچاردسون ۲۰ محاسبه شده است؟
آیا در نظریه سؤال - پاسخ ، نیکویی برازش انجام شده است ؟

پایایی:

آیا حداقل دو اندازه گیری وجود دارد؟
آیا فاصله بین (دو اندازه گیری) گزارش شده است؟
آیا بیماران در فاصله میان دو اندازه گیری، وضعیت پایداری داشته اند؟
آیا شرایط (نحوه اجرا، محیط، دستورالعملها) در هر دو اندازه گیری یکسان بوده است؟
آیا ضریب همبستگی درون طبقه ای در نمرات پیوسته، محاسبه شده است؟
آیا ضریب کاپا در نمرات اسمی، فاصله ای و دوازدهی محاسبه شده است؟
آیا در نمرات ترتیبی، کاپای وزنی محاسبه شده است؟

خطای اندازه گیری:

آیا حداقل دو روش سنجش وجود دارد؟
آیا بیماران در فاصله ی میان دو اندازه گیری، وضعیت پایداری داشته اند؟
آیا شرایط (نحوه اجرا، محیط، دستورالعملها) در هر دو اندازه گیری یکسان بوده است؟
آیا در نظریه های کلاسیک آزمون، خطای معیار اندازه گیری، حداقل تغییر قابل تشخیص یا محدوده ی توافق، محاسبه شده است؟

دوایی

آیا در روش طولی، حداقل دو ابزار استفاده شده است؟
 آیا زمان فاصله گذاری (بین دو آزمون) گزارش شده است؟
 آیا اتفاقاتی که در فاصله زمانی بین دو آزمون روی داده (به طور مثال، مداخله یا دیگر حوادث مرتبط) به طور کامل توضیح داده شده است؟

آیا نسبت بیماران (مثلا بر اساس میزان بهبودی یا وخامت بیماری) تغییر کرده است؟
الزامات طراحی برای آزمون فرضیه: زمانی که استاندارد طلایی موجود نباشد:
 آیا فرضیه ای مبنی بر تغییر نمرات، پیش از جمع آوری داده ها وجود داشته است؟
 آیا فرضیه، جهت، وجود ارتباط یا تفاوت میانگین های مورد انتظار در نمرات تغییر یافته ابزار وجود دارد؟
 آیا فرضیه میزان مورد انتظار مطلق یا نسبی همبستگی یا تفاوت میانگین ها را در نمرات تغییر یافته ابزار شامل شده است؟
 آیا توصیفی کامل از ابزارهای معیار، ارائه شده است؟
الزامات طراحی برای آزمون فرضیه: زمانی که استاندارد طلایی موجود باشد:
 آیا ملاک به کار برده شده برای تغییرات، می تواند به عنوان یک استاندارد طلایی منطقی، محسوب گردد؟
 آیا برای اعداد پیوسته، همبستگی بین نمرات تغییر یافته و سطح زیر منحنی عملیاتی دریافت کننده، منحنی راک، محاسبه می شود؟
 آیا برای ابزارهای دو ارزشی، حساسیت و ویژگی (تغییر یافته ها در مقابل تغییر نیافته ها) تعیین شده است؟

آیا پراکندگی کل نمرات در نمونه پژوهش، توصیف شده است؟
 آیا درصد نمره پاسخ دهنده ای که پایین ترین نمره کل ممکن را داشت، توصیف شده است؟
 آیا درصد نمره پاسخ دهنده ای که بالاترین نمره کل ممکن را داشت، توصیف شده است؟
 آیا نمرات و تغییر نمرات (به طور مثال میانگین) برای گروه ها و زیر گروه های مرتبط (بیماران یا عموم مردم)، ارائه شده است؟
 آیا حداقل تغییر با اهمیت یا حداقل تفاوت با اهمیت، مشخص شده است؟

Evidence Based Care Journal

<http://ebcj.mums.ac.ir/>

Psychometric Properties of Quality of Life Assessment Tools in Morbid Obesity: A Review of Literature

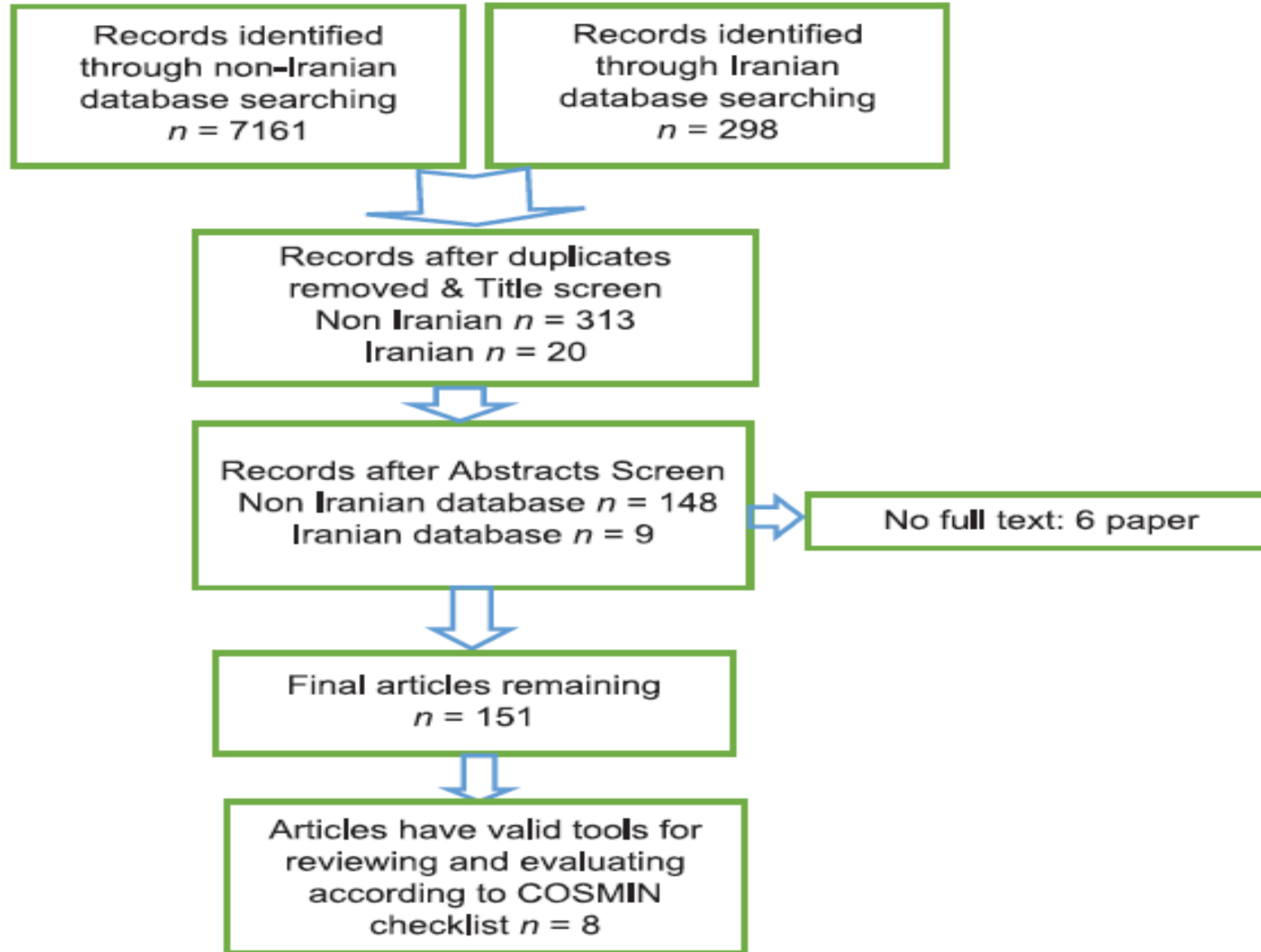
Negar Yazdani, Farkhondeh Sharif, Nasrin Elahi, Abbas Ebadi, Seyed Vahid Hosseini

Access this article online
Quick Response Code:

Website: www.jehp.net
DOI: 10.4103/jehp.jehp_391_20

Assessing sexual and reproductive health dimensions tools in women with type 1 diabetes mellitus with regard to Consensus-based Standards for the selection of health status Measurement Instruments checklist

Nasrin Azimi^{1,2}, Abbas Ebadi^{3,4}, Hamid Alavi Majd⁵, Assadollah Rajab⁶, Giti Ozgoli^{7,8}



Flow Diagram 1: Search procedure for studies relating to tools assessing sexual

Table 3

Tools assessing sexual and reproductive health dimension in women with Type 1 diabetes mellitus with regard to the Consensus-based Standards for the selection of health status Measurement Instruments Checklist

Tool	Country	Age group	Dimensions	Number of items	Content validity	Face validity	Construct validity			Reliability		Internal consistency	Responsiveness
							Cross-cultural validity	Structural validity	Hypothesis test	Reliability	Measurement error		
RHAB[18]	Russia and USA 2006	16-22 years	10 3 models Achieve normal BS Obtain prepregnancy counseling contraception	42	Based on pregnancy and diabetes Interview schedule[22]	No information	No information	Factor analysis	No information	No information		$\alpha=0.65-0.83$	No information
FSDS[23]	USA 2002	40±0.06 years	1	12	No information	No information	Yes[24]	Factor analysis	No information	Test re test $r \geq 0.91$	No information	$\alpha=0.86$	No information
FSFI[25]	USA 2000	Case: 21-69 years Control: 29-69 years	6 Desire - Arousal - Lubrication Orgasm	19	Yes	By expert panel	Yes[26]	Factor analysis Discriminant validity Divergent validity (marital adjustment test)	No information[28]	Test-retest 0.88	No information	$\alpha \geq 0.82$	No information[28]

Table 2. Tools measuring quality of life in morbid obesity patients

Test	Country of origin	Scaling	Domains	Number of items	Validity			Reliability		Interpretability	Responsiveness
					V1	V2	V3	R1	R2		
IWQOL-Lite	USA	Five-point Likert Scale: (0:never true to 4:always true)	Physical function (11 items), Self-esteem (7 items), Sexual life (4 items), Public distress (5 items), Work (4 items)	31	-	*	√**	√	√	-	-
OP Scale	Sweden	Four-point Likert scale: (4: Definitely bothered to 1: Definitely not bothered)	Psychosocial functioning : Private gatherings in my own home, in a friend's or relative's home, going to a restaurant, going to community activities, vacations away from home, trying on and buying clothes, bathing in public places, intimate	8	-	-	√	-	√	-	√

IWQOL: Impact of Weight on Quality of Life; OP: Obesity-Related Problem; BAROS: Bariatric Analysis and Reporting Outcome System; BOSS: Bariatric and Obesity-Specific Survey; M-A QOL: Moorehead-Ardelt Quality of Life; QOL OD: Quality Of Life, Obesity and Dietetics; BQL: The Bariatric Quality of Life. V1 - Content Validity; V2 - Criterion Validity; V3 - Construct Validity; R1 – Stability; R2 - Internal Consistency.

*Dash (-): It means that the property is not checked in the tool

**Checkmark (√): It means that the property is checked and is properly evaluated in the tool.

جدول شماره 6. رویکرد اصلاح شده GRADE برای درجه بندی کیفیت شواهد

کیفیت شواهد	پایین تر است اگر
بالا	<ul style="list-style-type: none"> ● خطر تورش
متوسط	<ol style="list-style-type: none"> 1. جدی 2. خیلی جدی
پایین	<ol style="list-style-type: none"> 3. ناسازگاری خیلی جدی
خیلی پایین	<ul style="list-style-type: none"> ● عدم ثبات 1. جدی 2. تحقق خیلی جدی <ul style="list-style-type: none"> ● دقت پایین 1. حجم نمونه 50-100 = 2. حجم نمونه < 50 <ul style="list-style-type: none"> ● پیچیدگی 1. جدی 2. خیلی جدی

Quality criteria were proposed for measurement properties of health status questionnaires

Caroline B. Terwee^{a,*}, Sandra D.M. Bot^a, Michael R. de Boer^{a,b},
Daniëlle A.W.M. van der Windt^{a,c}, Dirk L. Knol^{a,d}, Joost Dekker^{a,e},
Lex M. Bouter^a, Henrica C.W. de Vet^a

^a*EMGO Institute, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands*

^b*Department of Ophthalmology, VU University Medical Center, Amsterdam, The Netherlands*

^c*Department of General Practice, VU University Medical Center, Amsterdam, The Netherlands*

^d*Department of Clinical Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands*

^e*Department of Rehabilitation Medicine, VU University Medical Center, Amsterdam, The Netherlands*

GH.Joveini, OT Ph.D candidate, IUMS

Accepted 29 March 2006

Quality criteria for measurement properties of health status questionnaires

Property	Definition	Quality criteria ^{ab}
1. Content validity	The extent to which the domain of interest is comprehensively sampled by the items in the questionnaire	+A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection; ?A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method; -No target population involvement; 0No information found on target population involvement.
2. Internal consistency	The extent to which items in a (sub)scale are intercorrelated, thus measuring the same construct	+Factor analyses performed on adequate sample size ($7 * \# \text{ items}$ and ≥ 100) AND Cronbach's alpha(s) calculated per dimension AND Cronbach's alpha(s) between 0.70 and 0.95; ?No factor analysis OR doubtful design or method; -Cronbach's alpha(s) < 0.70 or > 0.95 , despite adequate design and method; 0No information found on internal consistency.
3. Criterion validity	The extent to which scores on a particular questionnaire relate to a gold standard	+Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥ 0.70 ; ?No convincing arguments that gold standard is "gold" OR doubtful design or method; -Correlation with gold standard < 0.70 , despite adequate design and method; 0No information found on criterion validity.

اهداف طراحی چک لیست “CosMIN”

► تعیین ویژگی های روانسجی که برای هر ابزار باید مورد بررسی قرار گیرد و ارائه تعریفی از این ویژگی ها

► ارائه معیارهایی جهت بررسی مناسب بودن این ویژگی ها

► تعیین استاندارد هایی جهت ارزیابی این ویژگی ها در طراحی مطالعات و تحلیل های آماری

Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist

**Caroline B. Terwee • Lidwine B. Mokkink •
Dirk L. Knol • Raymond W. J. G. Ostelo •
Lex M. Bouter • Henrica C. W. de Vet**

Box A. Internal consistency

		excellent	good	fair	poor
1	Does the scale consist of effect indicators, i.e. is it based on a reflective model? <i>Design requirements</i>				
2	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4	Was the sample size included in the internal consistency analysis adequate?	Adequate sample size (≥ 100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (< 30)
5	Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	Factor analysis performed in the study population	Authors refer to another study in which factor analysis was performed in a similar study population	Authors refer to another study in which factor analysis was performed, but not in a similar study population	Factor analysis NOT performed and no reference to another study
6	Was the sample size included in the unidimensionality analysis adequate?	7* #items and ≥ 100	5* #items and ≥ 100 OR 6-7* #items but < 100	5* #items but < 100	$< 5^*$ #items

سیستم ارزیابی چهار نقطه ای

برای هر مطالعه، یک قضاوت کلی روی کیفیت مطالعه خاص مورد نیاز است. بنابراین یک سیستم ارزیابی چهار نقطه ای براساس استانداردهای درون جعبه کاسمین توسعه یافت که می تواند به صورت **بسیار خوب، مناسب، مشکوک یا نامناسب** باشد. امتیاز کلی کیفیت هر مطالعه با در نظر گرفتن پایین ترین امتیاز هر استاندارد در جعبه تعیین می شود (یعنی "اصل بدترین نمره"). این امتیاز کلی در کیفیت مطالعات می تواند در درجه بندی کیفیت شواهد مورد استفاده قرار گیرد. زیرا نتایج مطالعات با کیفیت نامناسب، اعتماد را در یافته ها و نتیجه گیری کلی درباره ویژگی اندازه گیری PROM کاهش می دهد.

هر جعبه شامل یک سؤال استاندارد است که آیا ایرادهای روش شناسی مهم دیگری وجود دارد که بوسیله چک لیست پوشش داده نشود که خود منجر به یافته ها و نتایج همراه با سوگرایی گردد. برای مثال، در یک مطالعه پایایی، مداخلات باید مستقل باشد. مداخلات این مفهوم را می رساند که مداخله اولیه تحت تاثیر مداخله ثانویه نیست، به عنوان مثال در مداخله ثانویه بیمار نباید از نمرات اولین مداخله آگاه باشد (سوگرایی به خاطر آوردن).

RESEARCH ARTICLE

Evaluating the Psychometric Quality of Social Skills Measures: A Systematic Review

Reinie Cordier¹*, Renée Speyer², Yu-Wei Chen³, Sarah Wilkes-Gillan⁴, Ted Brown⁵‡, Helen Bourke-Taylor⁶‡, Kenji Doma²‡, Anthony Leicht²‡

1 School of Occupational Therapy and Social Work, Curtin University, Perth, WA, Australia, **2** College of Healthcare Sciences, James Cook University, Townsville, QLD, Australia, **3** Faculty of Health Sciences, The University of Sydney, Sydney, NSW, Australia, **4** School of Allied Health, Australian Catholic University, North Sydney, NSW, Australia, **5** Department of Occupational Therapy, Faculty of Medicine, Nursing and Health Sciences, Monash University—Peninsula Campus, Frankston, VIC, Australia, **6** School of Allied Health, Australian Catholic University, St Patricks Campus, Fitzroy, VIC, Australia

☉ These authors contributed equally to this work.

‡ These authors also contributed equally to this work.

* reinie.cordier@curtin.edu.au

so severe that it inhibits differentiation between more subtle psychometric qualities of instruments [63], a revised scoring was introduced. The outcome was presented as percentage of rating (Poor = 0–25.0%, Fair = 25.1%-50.0%, Good = 50.1%-75.0%, Excellent = 75.1%-100.0%). Given that some COSMIN items only have excellent and good as an option for rating, we calculated the total score for each psychometric property using the following formula to most accurately capture the quality of the psychometric properties:

$$\text{Total score for psychometric property} = \frac{(\text{Total score obtained} - \text{minimum score possible})}{(\text{Max score possible} - \text{minimum score possible})} \times 100$$

Table 6. Overview of the psychometric properties and methodological quality of occupational performance instruments.

Instrument & Author(s)	Year	Internal consistency*	Reliability*	Measurement error*	Content validity*	Structural validity*	Hypotheses testing*	Cross-cultural validity*	Criterion validity*
MMD									
Ricon et al. [8]	2013 ^A	Fair (42.9)	NR	NR	NR	NR	Fair (29.2)	NR	NR
PEGS									
Vroland-Nordstrom & Krumlinde-Sundholm [43, 44]	2012 ^A	NR	Excellent (79.5)	NR	NR	NR	Good (56.5)	Excellent (80.5)	NR
CAPE									
King et al. [46]	2004 ^M	Good (61.9)	Good (70.7)	NR	Good (64.3)	NR	NR	NR	NR
King et al. [47]	2006 ^A	NR	NR	NR	NR	NR	Good (74.6)	NR	NR
Colón et al. [45]	2008 ^A	Excellent (81.0)	NR	NR	NR	NR	Excellent (86.3)	Good (51.5)	NR
Ullenhag et al. [51]	2012 ^A	Excellent (76.2)	NR	NR	Excellent (78.6)	NR	Good (64.8)	Good (54.4)	NR

کاربردهای چک لیست نقد “CosMIN”

Systematic reviews of measurement properties

Authors of systematic reviews of measurement properties can use the COSMIN checklist to evaluate the **methodological quality of the included studies on measurement properties**. The assessment of the methodological quality of the included studies is an important step in any kind of systematic review because **low quality studies have a high risk of biased results**. The COSMIN checklist is comparable to similar checklists that have been developed for use in systematic reviews of other types of studies, such as the Delphi list for assessing the methodological quality of randomized clinical trials [9], the **QUADAS** checklist for assessing the methodological quality of diagnostic studies [10], and the **QUIPS** checklist for assessing the methodological quality of prognostic studies (available from JA Hayden, Toronto, Canada: jhayden@dal.ca).

Measurement instrument selection

Researchers who are selecting a measurement instrument for their study can use the COSMIN checklist to assess the quality of the available evidence on the measurement properties of the selected instrument(s) or of different available measurement instruments. They can use the COSMIN checklist in combination with criteria for good measurement properties (e.g. those developed by Terwee et al. [11]) to selected the best measurement instrument for a given purpose.

Identification of the need for further research on measurement properties

Application of the COSMIN checklist and taxonomy can also identify the need for further research on the measurement properties of a measurement instrument. The COSMIN taxonomy can be used to see whether all measurement properties have been evaluated. The taxonomy includes all measurement properties that should be evaluated when an instrument is used for evaluative purposes. The COSMIN checklist can be used to assess whether the available evidence on the measurement properties is of high quality.

Designing a study on measurement properties

Researchers who are designing a study on the measurement properties of a particular measurement instrument can use the COSMIN checklist to make sure that their study meets the standards for excellent quality. For example, a researcher using the COSMIN checklist to design a study on the construct validity of a PRO instrument may decide, based on the items in the COSMIN checklist box on hypothesis testing, to formulate and test specific hypotheses about expected mean differences between groups or expected correlations between the scores on the instrument of interest and other, related instruments. This will ascertain the quality of the validation study.

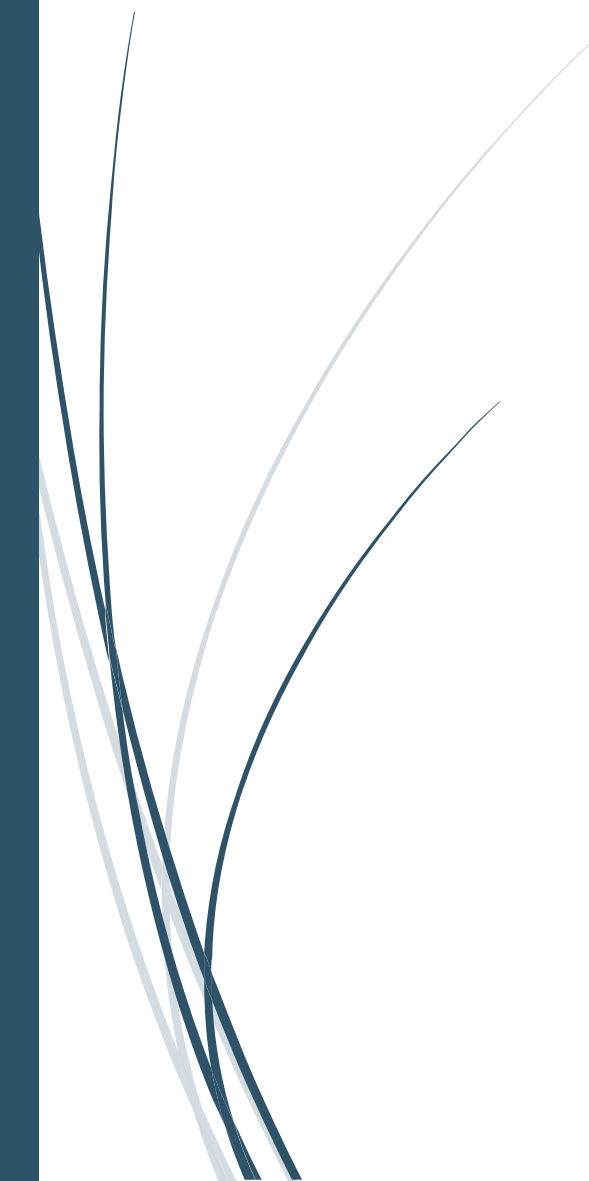
Reporting a study on measurement properties

Researchers who are reporting a study on measurement properties can use the COSMIN checklist to make sure that they report all information that is needed to enable an appropriate evaluation of the quality of their study. We recommend to use the COSMIN terminology and definitions of measurement properties to facilitate uniform reporting and avoid confusion in the literature on terms and definitions.

Reviewing the quality of a submitted manuscript on measurement properties

Editors or reviewers of submitted manuscripts can use the COSMIN checklist to assess whether the quality of a study on measurement properties is high enough to justify publication of the study. In addition, they can use the COSMIN checklist to identify issues that have not (yet) been (properly) reported. In the review process, the COSMIN checklist can be a useful tool to increase the quality of reporting of studies on measurement properties.

چند ایده برای نوشتن مقاله با کاسمین



PROTOCOL

Open Access



A systematic review protocol to evaluate the psychometric properties of measures of function within adult neuro-rehabilitation

Shannon Pike^{1,2*}, Natasha Anne Lannin^{1,3}, Anne Cusick^{4,5}, Kylie Wales^{6,7}, Lynne Turner-Stokes^{8,9} and Stephen Ashford^{8,9}



OTI

OCCUPATIONAL
THERAPY
INTERNATIONAL

[Explore this journal >](#)

Research Article

Survey Instruments for Knowledge, Skills, Attitudes and Behaviour Related to Evidence-based Practice in Occupational Therapy: A Systematic Review

[Helen Buchanan](#) , [Nandi Siegfried](#), [Jennifer Jelsma](#)

First published: 3 July 2015 [Full publication history](#)

DOI: [10.1002/oti.1398](https://doi.org/10.1002/oti.1398) [View/save citation](#)

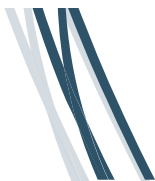
Cited by (CrossRef): 2 articles  [Check for updates](#) |  [Citation tools](#) ▼

PROTOCOL

Open Access




Functional assessments used by occupational therapists with older adults at risk of activity and participation limitations: a systematic review and evaluation of measurement properties

Kylie Wales^{1*}, Lindy Clemson¹, Natasha A Lannin² and Ian D Cameron³




RESEARCH ARTICLE

Child-Report Measures of Occupational Performance: A Systematic Review

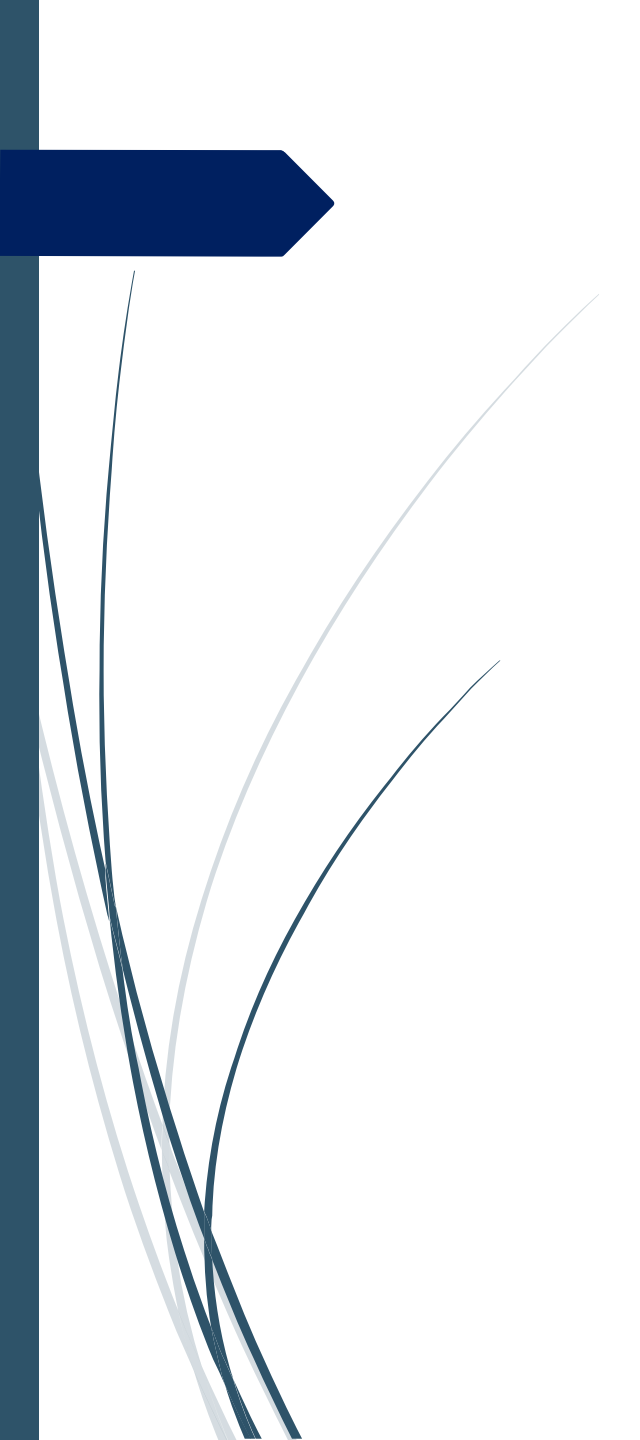
Reinie Cordier¹^{*}, Yu-Wei Chen²[Ⓞ], Renée Speyer³[Ⓞ], Rebekah Totino¹[‡], Kenji Doma³[‡], Anthony Leicht³[‡], Nicole Brown³[‡], Belinda Cuomo¹[‡]

1 School of Occupational Therapy and Social Work, Curtin University, Perth, WA, Australia, 2 Faculty of Health Sciences, The University of Sydney, Sydney, NSW, Australia, 3 College of Healthcare Sciences, James Cook University, Townsville, QLD, Australia

 These authors contributed equally to this work.

[‡] These authors also contributed equally to this work.

^{*} reinie.cordier@curtin.edu.au



OUTCOME MEASURES OF PHYSICAL
FUNCTION IN ADULTS with a UNILATERAL
LOWER LIMB AMPUTATION
DURING PROSTHETIC REHABILITATION:
USE IN CLINICAL PRACTICE AND
PSYCHOMETRIC PROPERTIES

JUDY SCOPES

A thesis submitted in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy

QUEEN MARGARET UNIVERSITY

2016



Tools for Observational Gait Analysis in Patients With Stroke: A Systematic Review

Francesco Ferrarello, Valeria Anna Maria Bianchi, [Marco Baccini](#), Gaia Rubbieri, [Enrico Mossello](#), Maria Chiara Cavallini, Niccolò Marchionni, [Mauro Di Bari](#)

Background. Stroke severely affects walking ability, and assessment of gait kinematics is important in defining diagnosis, planning treatment, and evaluating interventions in stroke rehabilitation. Although observational gait analysis is the most common approach to evaluate gait kinematics, tools useful for this purpose have received little attention in the scientific literature and have not been thoroughly reviewed.

Objectives. The aims of this systematic review were to identify tools proposed to conduct observational gait analysis in adults with a stroke, to summarize evidence concerning their quality, and to assess their implementation in rehabilitation research and clinical practice.

Methods. An extensive search was performed of original articles reporting on visual/observational tools developed to investigate gait kinematics in adults with a stroke. Two reviewers independently selected studies, extracted data, assessed quality of the included studies, and scored the metric properties and clinical utility of each tool. Rigor in reporting metric properties and dissemination of the tools also was evaluated.

Results. Five tools were identified, not all of which had been tested adequately for their metric properties. Evaluation of content validity was partially satisfactory. Reliability was poorly investigated in all but one tool. Concurrent validity and sensitivity to change were shown for 3 and 2 tools, respectively. Overall, adequate levels of quality were rarely reached. The dissemination of the tools was poor.

Conclusions. Based on critical appraisal, the Gait Assessment and Intervention Tool shows a good level of quality, and its use in stroke rehabilitation is recommended. Rigorous studies are needed for the other tools in order to establish their usefulness.

نتیجه گیری

آشنایی با این چک لیست می تواند در توانمند کردن محققین و درمانگران در انتخاب ابزار مناسب جهت ارزیابی ویژگی های مورد نظر در افراد بسیار تاثیر گذار باشد.