

## Review Paper

# The Framework of Designing, Reporting, and Evaluating the Methodological Quality for Psychometric Studies: The Revised COSMIN Checklists



Mohammad Zare<sup>1</sup>, Zahra Tagharrobi<sup>1\*</sup>

1. Trauma Nursing Research Center, Faculty of Nursing and Midwifery, Kashan University of Medical Sciences, Kashan, Iran.



**Citation** Zare M, Tagharrobi Z. [The Framework of Designing, Reporting, and Evaluating the Methodological Quality for Psychometric Studies: The Revised COSMIN Checklists (Persian)]. *Qom Univ Med Sci J.* 2023; 16(10):786-803. <https://doi.org/10.32598/qums.16.10.2751.1>

<https://doi.org/10.32598/qums.16.10.2751.1>



Received: 16 Oct 2022

Accepted: 14 Nov 2022

Available Online: 01 Jan 2023

### Keywords:

Psychometrics, Checklist, Bias, Study design, Research report

## ABSTRACT

**Background and Objectives** One of the challenges in studies related to measuring patient-reported outcomes (PRO) is the selection of an appropriate measurement tool. In this regard, compliance with standards in design and reporting of psychometric studies is necessity. In addition, researchers should be able to correctly evaluate the methodological quality of psychometric studies. Considering the presentation of the revised forms for the consensus-based standards for the selection of health Measurement instruments (COSMIN) checklist, this study aims to review these checklists.

**Methods** In this narrative review study, a search was conducted in national and international databases for articles measured the quality of psychometric studies by COSMIN checklist using related keywords in Persian and English with no time limitations. The results are presented in two sections, including the history of COSMIN checklist and the introduction of its revised forms.

**Results** There are four COSMIN checklists, including the COSMIN Risk of Bias Checklist, the COSMIN Risk of Bias Tool (by focusing on reliability and measurement error), the COSMIN Study Design Checklist, and the COSMIN Reporting Guideline.

**Conclusion** Considering the increasing number of psychometric studies on the PRO measurement tools, the COSMIN checklists can provide comprehensive indicators for designing, reporting, and evaluating these studies.

### \* Corresponding Author:

Zahra Tagharrobi

Address: Trauma Nursing Research Center, Faculty of Nursing and Midwifery, Kashan University of Medical Sciences, Kashan, Iran.

Tel: +98 (31) 55103119

E-Mail: tagharrobi\_z@kaums.ac.ir



## Extended Abstract

### Introduction

One of the challenges in studies related to measuring patient-reported outcomes (PRO) is the selection of appropriate measurement tools. Psychometric studies introduce reliable tools. Considering the increasing number of instruments for PROs, it is necessary to systematically study the psychometric properties of the instruments to select and use the best tools. Existence of a comprehensive and valid scale or checklist to evaluate the quality of psychometric studies is necessary. One of these tools is COSMIN (Consensus-based Standards for the selection of health Measurement instruments) checklist. Considering the presentation of the revised forms for this checklist and the importance of evaluating and reporting psychometric instruments, this study aims to review the revised forms of COSMIN checklist.

### Methods

In this narrative review study, a search was conducted in international databases ([PubMed](#), [Scopus](#), [Proquest](#), [Web of Science](#), and [CINAHL](#)), and national databases ([Magiran](#), [IranDoc](#), [ISC](#), and [SID](#)) for all articles that used COSMIN checklist to measure the quality of psychometric studies using related keywords in Persian and English without considering time limitations

### Results

In the first study using the Delphi method (2006), a working group consisting of 57 experts in epidemiology, measurement, statistics, and biostatistics, introduced a classification of characteristics for instruments in three areas of validity (including content and face validity, structural validity, hypothesis testing, cross-cultural validity, and criterion validity), reliability (including internal consistency, reliability, and measurement error) and responsiveness (including responsiveness and interpretability). Considering the limitations of the initial version, the COSMIN committee changed it to strengthen its reliability and ease of use. These changes led to creating four versions including (1) the COSMIN Risk of Bias Checklist, (2) the COSMIN Risk of Bias Tool to assess the quality of studies by focusing on reliability and measurement error, (3) the COSMIN Study Design Checklist, and (4) the COSMIN Reporting Guideline. With the introduction of new checklists, the COSMIN committee advised the researchers not to use the original version and only use the newly introduced checklists according to the study objectives.

### Discussion

Considering the need to use valid and reliable PRO measures in research and clinic settings, as well as the increasing number of instruments for PRO constructs, there was a need for a scale to provide criteria for measuring the psychometric properties of the instruments. In this regard, the COSMIN checklists were introduced. However, they have some limitations such as being specialized and the need for teaching how to complete it, being opinion-based (not evidence-based), lack of sufficient evidence to confirm the reliability of the checklist, failure to determine the importance of some psychometric properties compared to others, and lack of assessing essential details in evaluation of some psychometric properties. In overall, COSMIN checklists can provide logical indicators and standards for measuring the psychometric properties of instruments; therefore, they can be used for measuring the quality of psychometric studies, designing studies, or reporting psychometric studies by researchers, reviewers, and journal editors.

### Ethical Considerations

#### Compliance with ethical guidelines

This study was approved by the Ethics Committee of [Kashan University of Medical Sciences](#) (Code: IR.KAUMS.NUHEPM.REC.1401.047).

#### Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

#### Authors contributions

All authors equally contributed to preparing this article.

#### Conflicts of interest

The authors declared no conflict of interest.

#### Acknowledgements

In this way, all the researchers, especially COSMIN researchers, whose articles were reviewed in this study, as well as the Ethics Committee of [Kashan University of Medical Sciences](#) and Healthcare Services, are thanked and appreciated.

This Page Intentionally Left Blank

## مقاله مروری

## چارچوب طراحی، گزارش و ارزیابی کیفیت روش‌شناسی مطالعات روان‌سنجی: چک‌لیست‌های تجدیدنظرشده کاسمین

محمد زارع<sup>۱</sup>، زهرا تقریبی<sup>۱</sup>

۱. مرکز تحقیقات پرستاری تروما، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی کاشان، کاشان، ایران

Use your device to scan  
and read the article online

**Citation** Zare M, Tagharrobi Z. [The Framework of Designing, Reporting, and Evaluating the Methodological Quality for Psychometric Studies: The Revised COSMIN Checklists (Persian)]. *Qom Univ Med Sci J.* 2023; 16(10):786-803. <https://doi.org/10.32598/qums.16.10.2751.1>

<https://doi.org/10.32598/qums.16.10.2751.1>

## چکیده

**زمینه و هدف:** یکی از چالش‌های مطرح در حوزه بالین و پژوهش در ارتباط با سنجش پیامدهای مرتبط با سلامت، انتخاب ابزار اندازه‌گیری مناسب است. در این راستا رعایت استانداردها در طراحی و گزارش مطالعات روان‌سنجی ابزار یک ضرورت است؛ به‌علاوه پژوهشگران باید بتوانند به شکل صحیح و اصولی کیفیت روش‌شناسی مطالعات مربوط به روان‌سنجی و معرفی ابزارها را ارزیابی و نقد کنند. پیش از این، مقاله‌ای به زبان فارسی به معرفی شاخص‌های انتخاب ابزارهای اندازه‌گیری پرداخته بود، اما با توجه به ارائه فرم تجدیدنظرشده برای این چک‌لیست و اهمیت طراحی، گزارش و نقد مقالات روان‌سنجی ابزار، مقاله حاضر با هدف معرفی و توصیف چک‌لیست‌های تجدیدنظرشده کاسمین تنظیم شده است.

**روش بررسی:** این مطالعه از نوع مرور روایتی بود. مستندات علمی پیرامون سنجش کیفیت مطالعات روان‌سنجی، توسط چک‌لیست‌های کاسمین با استفاده از کلیدواژه‌های Psychometric و Instrument و معادل آن‌ها و بدون در نظر گرفتن محدودیت زمانی در بانک‌های اطلاعاتی ملی و بین‌المللی بررسی شد.

**یافته‌ها:** نتایج این مطالعه در ۲ بخش شامل تاریخچه چک‌لیست کاسمین و چک‌لیست‌های معرفی‌شده، ارائه شده است. تاکنون، ۴ چک‌لیست کاسمین شامل چک‌لیست‌های خطر سوگیری، خطر سوگیری جهت ارزیابی کیفیت مطالعات با تأکید بر پایایی یا خطای اندازه‌گیری، طراحی مطالعه و راهنمای گزارش، معرفی شده و در دسترس قرار گرفته است.

**نتیجه‌گیری:** با توجه به نیاز به استفاده از ابزارهای معتبر و قابل‌اطمینان و همچنین تعدد روزافزون مطالعات روان‌سنجی ابزارهای سنجش پیامدهای مرتبط با سلامت، چک‌لیست‌های کاسمین می‌توانند با ارائه شاخص‌های جامع و منطقی در حیطه‌های طراحی، گزارش و ارزیابی کیفیت روش‌شناسی مطالعات، این مهم را برای پژوهشگران تسهیل کنند.

تاریخ دریافت: ۲۴ مهر ۱۴۰۱

تاریخ پذیرش: ۲۳ آبان ۱۴۰۱

تاریخ انتشار: ۱۱ دی ۱۴۰۱

## کلیدواژه‌ها:

روان‌سنجی، چک‌لیست، سوگیری، طراحی پژوهش، گزارش پژوهش

## \* نویسنده مسئول:

زهرا تقریبی

نشانی: کاشان، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی کاشان، دانشکده پرستاری و مامایی، مرکز تحقیقات پرستاری تروما.

تلفن: ۰۳۱۱۹۵۵۱۰۳۱۱ (۳۱) ۹۸+

رایانامه: tagharrobi\_z@kaums.ac.ir



## مقدمه

بررسی سیستماتیک کیفیت مطالعات روان‌سنجی ابزار با تمرکز بر ویژگی‌های اندازه‌گیری ابزارها استفاده کرد. همچنین، این چک‌لیست می‌تواند به‌عنوان راهنمایی برای طراحی یا گزارش مطالعات روان‌سنجی ابزار استفاده شود. علاوه‌براین، پژوهشگران، داوران و ویراستاران مجلات می‌توانند از این چک‌لیست جهت ارزیابی کیفیت ویژگی‌های روان‌سنجی مطالعات ابزارسازی استفاده کنند [۷، ۱۲]. اگرچه چک‌لیست کاسمین در ابتدا صرفاً برای ارزیابی سنج‌های پیامدی گزارش‌شده توسط بیمار توسعه یافت و تأیید شد، اما در برخی مطالعات از معیارهای آن جهت ارزیابی سنج‌های پیامدی گزارش‌شده توسط غیربیماران نیز استفاده شده است [۱۳].

باتوجه به تعدد ابزارهای معرفی‌شده و نیاز روزافزون به ارزیابی کیفیت مطالعات روان‌سنجی ابزار و همچنین تغییرات ساختاری چک‌لیست کاسمین در سال‌های اخیر، مقاله حاضر با هدف معرفی چک‌لیست‌های کاسمین و توصیف راهنمای گزارش کاسمین انجام شد.

## روش بررسی

مطالعه حاضر، یک مرور روایتی است. جامعه پژوهش شامل کلیه مقالات علمی در زمینه ارزیابی کیفیت مطالعات روان‌سنجی ابزار با استفاده از چک‌لیست کاسمین است. در این مطالعه، اطلاعات با جست‌وجوی نظام‌مند در پایگاه‌های اطلاعاتی بین‌المللی وب‌آوساینس، پابمد، اسکوپوس، سیناهل، پروکوئست و همچنین پایگاه‌های فارسی‌زبان ایران‌داک، مگ‌ایران، مرکز اطلاعات علمی جهاد دانشگاهی و پایگاه استنادی علوم جهان اسلام بدون در نظر گرفتن محدودیت زمانی (تا جولای ۲۰۲۲) گردآوری شد.

جست‌وجوی مقالات مرتبط براساس استراتژی پیشنهادشده توسط پژوهشگران کاسمین [۱۴] و با استفاده از کلیدواژه‌های Instrument، method، compar، psychometric clinimetric، outcome، assess، measure، observ، reproducible، test یا معادل فارسی آن‌ها با استفاده از عملگرهای AND و OR انجام شد. جهت اطمینان از بازیابی کلیه منابع، جست‌وجو به ۲ روش نسلی و نسبی نیز انجام شد.

## یافته‌ها

نتایج این مطالعه در ۲ بخش شامل تاریخچه چک‌لیست کاسمین و معرفی چک‌لیست‌های کاسمین ارائه شده است.

## تاریخچه چک‌لیست کاسمین

در اولین مطالعه بین‌المللی کاسمین که در سال ۲۰۰۶ به روش دلفی انجام شد، کارگروهی متشکل از ۵۷ متخصص در

سنج‌های پیامدی گزارش‌شده توسط بیمار<sup>۱</sup> شامل جنبه‌های مختلفی از وضعیت سلامتی هستند که مستقیماً توسط بیماران بیان می‌شوند و سایر افراد در تفسیر آن دخالت ندارند [۱]. یکی از چالش‌های ارزیابی این سنج‌ها، کیفیت ابزارهای اندازه‌گیری است [۲، ۳]. در این زمینه، مطالعات روان‌سنجی وظیفه سنجش ویژگی‌های هر یک از ابزارهای معرفی‌شده را برعهده دارند؛ به‌طوری که مطالعات ارائه‌دهنده ویژگی‌های روان‌سنجی مطلوب، ابزار قابل‌اعتمادتری را معرفی می‌کنند [۴].

باتوجه به تعدد روزافزون ابزارهای اندازه‌گیری سازه‌های مرتبط با سلامتی (به‌ویژه سنج‌های پیامدی گزارش‌شده توسط بیمار)، بررسی سیستماتیک ویژگی‌های روان‌سنجی ابزارها جهت انتخاب و کاربست بهترین ابزار ضروری است [۵]. اگر کیفیت سنجش ویژگی‌های روان‌سنجی یک ابزار مطلوب باشد، می‌توان از نتایج در راستای کیفیت مطلوب ابزار موردنظر استفاده کرد؛ از سوی دیگر، زمانی که کیفیت روش‌شناسی یک مطالعه روان‌سنجی ناکافی باشد، نمی‌توان به نتایج مطالعه اعتماد کرد و کیفیت ابزار موردمطالعه نامشخص باقی خواهد ماند [۶]. بنابراین وجود یک مقیاس یا چک‌لیست جهت ارزیابی کیفیت مطالعات روان‌سنجی ابزار، مهم تلقی می‌شود [۷، ۸].

برخی از پژوهشگران حوزه روان‌سنجی ابزار، کیفیت روش‌شناسی مطالعات را در قالب مطالعات مروری ارزیابی کرده‌اند [۹-۱۱]؛ برای مثال، هایود و همکاران براساس تعداد و نوع ارزیابی‌ها [۹]، ویند و همکاران مطابق هدف مطالعات، جامعه پژوهش، روش تجزیه و تحلیل و ارائه نتایج آماری [۱۰] و مارینوس و همکاران براساس مناسب بودن حجم نمونه و آنالیزها [۱۱] کیفیت مطالعات روان‌سنجی را بررسی کردند. باین‌حال، روش یکسانی جهت سنجش کیفیت این مطالعات، مورد استفاده قرار نگرفته است. اختلاف در معیارهای تصمیم‌گیری جهت تعیین کیفیت مطالعات می‌تواند به عدم قضاوت صحیح و سردرگمی بیش‌ازپیش متخصصان حوزه‌های مرتبط منجر شود. بنابراین وجود یک مقیاس جامع و معتبر لازم و ضروری است [۴].

به این منظور موکینک و همکاران، در یک مطالعه بین‌المللی براساس روش دلفی در سال ۲۰۰۶ با همکاری یک گروه چندرشته‌ای از محققان حوزه روان‌سنجی ابزار، استانداردهای مبتنی بر اجماع جهت انتخاب ابزارهای اندازه‌گیری سلامت<sup>۲</sup> را توسعه دادند. این مطالعه با تأکید بر ارزیابی کیفیت روش‌شناسی مطالعات روان‌سنجی انجام شد و نهایتاً منجر به معرفی چک‌لیست کاسمین شد [۴]. چک‌لیست کاسمین را می‌توان در

1. Patient-Reported Outcome Measures (PROMs)  
 2. COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN)

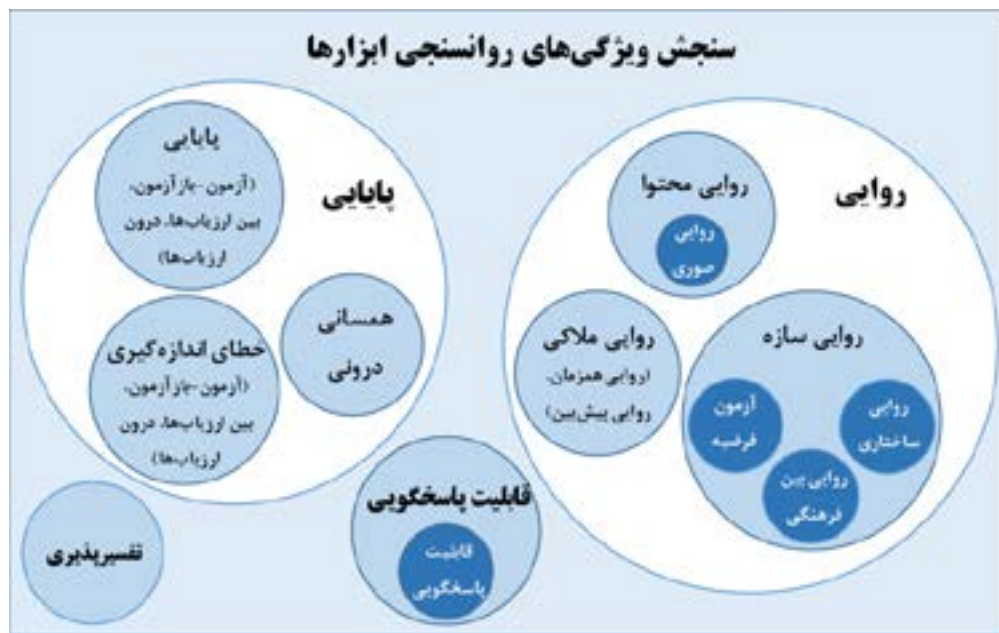
جدول ۱. تعاریف ابعاد، ویژگی‌ها و جنبه‌های اندازه‌گیری براساس طبقه‌بندی کاسمین

تعریف	عبارت	
	جنبه موردنظر	ویژگی اندازه‌گیری
درجه‌ای که در آن، اندازه‌گیری عاری از خطای اندازه‌گیری است.		پایایی (Reliability)
درجه‌ای از همبستگی متقابل بین گویه‌های ابزار را نشان می‌دهد.		همسانی درونی (Internal Consistency)
نسبت واریانس کل اندازه‌گیری‌ها که به تفاوت بین نمرات اشاره دارد.		پایایی
خطای سیستماتیک و تصادفی نمرات یک ابزار که به تغییرات واقعی در سازه موردنظر نسبت داده نمی‌شود.		خطای اندازه‌گیری (Measurement Error)
درجه‌ای که یک ابزار مرتبط با سلامت، بتواند سازه مورد نظر را بسنجد.		روایی (Validity)
درجه‌ای که محتوای یک ابزار، بتواند بازتاب مناسبی از سازه مورد اندازه‌گیری را ارائه دهد.		روایی محتوا (Content Validity)
درجه‌ای که ظاهر گویه‌های یک ابزار، انعکاس کافی از ساختار موردسنجش را ارائه دهد.	روایی صورتی (Face Validity)	
درجه‌ای که نمرات یک ابزار بتواند سازگاری سازه مورد اندازه‌گیری را با فرضیه‌های پژوهش (روابط داخلی، روابط با نمرات سایر ابزارها یا تفاوت بین گروه‌ها) نشان دهد.		روایی سازه (Construct Validity)
درجه‌ای از نمرات یک ابزار که بتواند بازتاب مناسبی از ابعاد ساختاری سازه موردسنجش را ارائه دهد.	روایی ساختاری (Structural Validity)	
مشابه روایی سازه	آزمون فرضیه (Hypothesis Testing)	
درجه‌ای از عملکرد گویه‌ها در یک ابزار ترجمه‌شده یا اقتباس‌شده که بازتاب مناسبی از عملکرد گویه‌های نسخه اصلی را ارائه دهد.	روایی بین‌فرهنگی (Cross-cultural Validity)	
درجه‌ای از نمرات یک ابزار که بازتاب مطلوبی از استاندارد طلایی سازه ارائه دهد.		روایی ملاکی (Criterion Validity)
توانایی یک ابزار در تشخیص تغییرات ایجادشده در طول زمان در سازه‌ای که مورد اندازه‌گیری قرار می‌گیرد.		قابلیت پاسخ‌گویی (Responsiveness)
مشابه قابلیت پاسخ‌گویی	قابلیت پاسخ‌گویی	
درجه‌ای از توانایی ابزار که قادر باشد امتیازات کمی را به معنای کیفی و قابل‌درک (نظیر اهمیت بالینی) سازه نسبت دهد.		تفسیرپذیری (Interpretability)

جداگانه ارزیابی می‌شود. بخش‌های چک‌لیست کاسمین عبارت‌اند از: ۹ بخش شامل سنجش ویژگی‌های روان‌سنجی (همسانی درونی، پایایی، خطای اندازه‌گیری، روایی محتوا، روایی ساختاری، آزمون فرضیه، روایی بین‌فرهنگی، روایی ملاکی و قابلیت پاسخ‌گویی)، ۱ بخش حاوی استانداردهای مطالعات درمورد تفسیرپذیری و ۲ بخش شامل الزامات کلی جهت ارزیابی تعمیم‌پذیری<sup>۳</sup> و ارزیابی مقالاتی که از نظریه سؤال-پاسخ استفاده کردند. تعداد سؤالات در هر بخش از چک‌لیست کاسمین از ۵

رشته‌هایی نظیر اپیدمیولوژی، سنجش و اندازه‌گیری، آمار و آمار زیستی (با میانگین ۲۰ سال سابقه در حوزه روان‌سنجی ابزار)، یک طبقه‌بندی از ویژگی‌های اندازه‌گیری ابزارها معرفی کردند که در تصویر شماره ۱ نشان داده شده است [۴، ۸، ۱۲]. براین اساس ۳ حوزه روایی، پایایی و قابلیت پاسخ‌گویی از یکدیگر تفکیک شدند که تعاریف هر یک از ویژگی‌ها مختصراً در جدول شماره ۱ بیان شده است [۸].

براین اساس، چک‌لیستی متشکل از دوازده بخش ارائه شد که هر بخش بر یک ویژگی خاص اندازه‌گیری، تمرکز دارد و به صورت



تصویر ۱. طبقه‌بندی ویژگی‌های روانسنجی ابزارهای سنجش‌های پیامدی گزارش‌شده توسط بیمار بر اساس طبقه‌بندی COSMIN<sup>۶</sup> دانشگاه علوم پزشکی قم

پس از مدتی باتوجه به محدودیت‌های نسخه اصلی چک‌لیست کاسمین و با هدف سهولت استفاده جهت اهداف مختلف، پژوهشگران کاسمین تصمیم به ویرایش چک‌لیست و ارائه نسخه‌های مجزا گرفتند [۱۸]؛ چراکه استانداردهای موجود در هر چک‌لیست کمی متفاوت خواهد بود. برای مثال، استاندارد تعیین حجم نمونه، هنگام طراحی مطالعات مفید و ضروری است و هنگام گزارش‌دهی مطالعات روان‌سنجی، بیشتر استانداردهای گزارش مقادیر گمشده مطرح است [۱۹].

#### چک‌لیست‌های کاسمین

پس از مرور گسترده و تشکیل کمیته‌های مجزا منطبق بر اهداف کاسمین، چک‌لیست‌های زیر از یکدیگر تفکیک و به‌صورت جداگانه معرفی شدند:

- چک‌لیست خطر سوگیری کاسمین<sup>۶</sup> [۱۸]؛
- ابزار خطر سوگیری کاسمین جهت ارزیابی کیفیت مطالعات با تأکید بر پایایی یا خطای اندازه‌گیری<sup>۷</sup> [۲۰]؛
- چک‌لیست طراحی مطالعه کاسمین<sup>۸</sup> [۲۱]؛
- راهنمای گزارش کاسمین<sup>۹</sup> [۱].

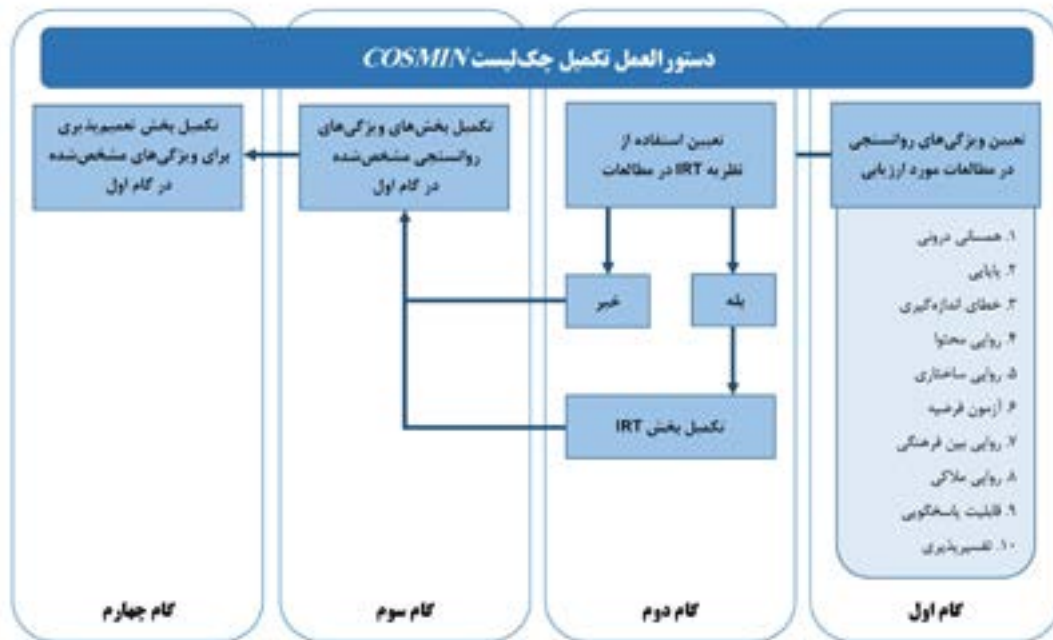
6. COSMIN Risk of Bias Checklist  
7. COSMIN Risk of Bias Tool to assess the quality of studies on reliability and measurement error  
8. COSMIN Study Design Checklist  
9. COSMIN Reporting Guideline

تا ۱۸ مورد، متغیر است و با گزینه‌های «+» (بله)، «-» (خیر)، «؟» (نامشخص) و یا «NA» (غیرقابل انجام) قابلیت پاسخ‌دهی دارند [۴، ۷، ۸].

نسخه اولیه چک‌لیست کاسمین در یک فرایند ۴ مرحله‌ای تکمیل می‌شود که در تصویر شماره ۲ نشان داده شده است [۷]. در ابتدا باید مشخص شود کدامیک از ویژگی‌های روان‌سنجی باید در مطالعه موردنظر، ارزیابی شوند. در مرحله دوم، باتوجه به تفاوت بین نظریه‌های کلاسیک آزمون<sup>۵</sup> و سؤال-پاسخ، روش‌های آماری مرتبط نیز بررسی می‌شوند [۱۵]. نظریه کلاسیک آزمون، عموماً بر عملکرد آزمون‌ها و نمرات کسب‌شده تمرکز دارد؛ درحالی‌که تمرکز نظریه سؤال-پاسخ بر عملکرد سؤالات و توانایی و نحوه برخورد پاسخ‌دهندگان به سؤال است [۱۶]. در مرحله سوم، باتوجه به نوع ارزیابی تعیین‌شده در مرحله اول، بخش‌های موردنظر برای مطالعه تکمیل می‌شود. به‌عنوان مرحله نهایی، تصمیم‌پذیری برای ویژگی‌های شناسایی‌شده در مرحله اول ارزیابی می‌شود [۱۵].

در نسخه اولیه چک‌لیست کاسمین، چگونگی نمره‌دهی کلی کیفیت مطالعات روان‌سنجی بیان نشده است. در واقع، هدف این چک‌لیست، تبدیل کیفیت به کمیت نیست، چراکه همه معیارهای کیفیت روان‌سنجی از اهمیت یکسانی برخوردار نیستند [۱۷]. بااین‌حال، روش‌های مختلفی در این زمینه بیان شده است. در نهایت پژوهشگران کاسمین در سال ۲۰۱۲، سیستم رتبه‌بندی ۴ امتیازی جهت ارزیابی کیفیت مطالعات روان‌سنجی را پیشنهاد دادند [۶].

5. Classical Test Theory (CTT)



تصویر ۲. فرایند تکمیل چکلیست کاسمین

هر مقیاس یا زیرمقیاس و ترجیح در کاربست روش آماری آلفای کرونباخ بود [۱۸]. مشابه نسخه اصلی چکلیست، در چکلیست خطر سوگیری کاسمین نیز روش‌های آماری مبتنی بر نظریه‌های کلاسیک آزمون و سؤال-پاسخ و رویکرد اندازه‌گیری راش گنجانده شده است.

این چکلیست از ۱۰ بخش استانداردهای توسعه سنج‌های پیامدی گزارش شده توسط بیمار تشکیل شده است که شامل [۱] توسعه سنج‌های پیامدی گزارش شده توسط بیمار (طراحی سنج پیامدی گزارش شده توسط بیمار و مصاحبه شناختی یا آزمون اولیه)، [۲] روایی محتوا (نظرسنجی در ارتباط با مرتبط بودن، جامعیت و قابل درک بودن از بیماران و متخصصان)، [۳] روایی ساختاری، [۴] همسانی درونی، [۵] روایی بین فرهنگی / ثبات اندازه‌گیری<sup>۱۱</sup>، [۶] پایایی، [۷] خطای اندازه‌گیری، [۸] روایی ملاکی، [۹] آزمون فرضیه جهت روایی سازه (روایی همگرا و روایی افتراقی یا مقایسه گروه‌های شناخته شده) و [۱۰] قابلیت پاسخ‌گویی (رویکرد معیاری شامل مقایسه با استاندارد طلایی و رویکرد سازه‌ای شامل مقایسه با سایر ابزارهای سنجش پیامد، مقایسه بین گروه‌های شناخته شده و مداخلات قبل و بعد جهت آزمون فرضیه) است [۱، ۱۸، ۲۳].

با معرفی چکلیست‌های جدید، هسته مرکزی کاسمین به پژوهشگران توصیه کردند ضمن عدم استفاده از نسخه اولیه، از این پس صرفاً از نسخه‌های جدید معرفی شده استفاده کنند [۱۹].

### چکلیست خطر سوگیری کاسمین

چکلیست خطر سوگیری کاسمین [۱۸] به‌عنوان اولین جایگزین نسخه اصلی [۸] در سال ۲۰۱۸ معرفی شد. این چکلیست با هدف ارزیابی کیفیت روش‌شناسی مطالعات روان‌سنجی و غربالگری انواع سوگیری‌ها ارائه شده است [۱۸]. عبارت سوگیری<sup>۱۰</sup>، براساس دستورالعمل کاکرین برای مرورهای سیستماتیک کارآزمایی‌ها و مطالعات تشخیصی استفاده شده است. سوگیری عبارت است از هرگونه مداخله، تعصب یا نفوذ در مطالعه که به تغییر در نتایج منجر شود. انواع سوگیری شامل سوگیری انتخاب، سوگیری اطلاعات و سوگیری مخدوش‌شدگی می‌شود [۲۲]. بررسی وجود برخی از انواع سوگیری در مطالعات روان‌سنجی ابزار، مهم و حیاتی هستند [۱۸].

چکلیست خطر سوگیری کاسمین شامل استانداردها و الزامات طراحی و روش‌های آماری مناسب در مطالعات روان‌سنجی ابزار است که برای هر ویژگی، یک بخش سوگیری نیز مطرح شده است [۲۳]. این چکلیست در قالب یک مطالعه دلفی با اجماع گروهی متشکل از ۵۷ متخصص روان‌سنجی طراحی شده است. برخی از تغییرات این نسخه شامل مرتب کردن ویژگی‌های سنجش روان‌سنجی ابزار، سنجش جداگانه همسانی درونی برای



## جدول ۲. راهنمای گزارش کاسمین (نسخه آگوست ۲۰۲۱)

توصیه‌های عمومی گزارش‌دهی مرتبط با کلیه مطالعات روان‌سنجی ابزار		
شماره گویه	نام گویه	توصیف گویه
T1	سنجه‌های پیامدی گزارش‌شده توسط بیمار	نام ابزار(های) سنجش پیامد خودگزارش‌دهی بیمار مورد مطالعه (و در صورت لزوم، نسخه ابزار) ذکر شود.
T2	ویژگی موردسنجش	کدام شاخص یا ویژگی‌ها مورد مطالعه قرار گرفته است (به‌عنوان مثال، اگر تعداد زیادی از ویژگی‌ها بررسی شده‌اند).
T3	نمونه پژوهش	در صورت لزوم، کلیات ویژگی‌های نمونه مورد مطالعه (نظیر شرایط و زبان) و همچنین هرگونه مداخله یا مواجهه (نظیر درمان) توصیف شود.
A1	سنجه‌های پیامدی گزارش‌شده توسط بیمار	نام ابزار(های) در حال مطالعه (در صورت لزوم، همراه با نسخه و زبان؛ به‌طور مثال SF-۳۶ یا SF-۱۲)، ارتباط با بانک گویه (نظیر ابزارهای سنجه‌های پیامدی گزارش‌شده توسط بیمار) و همچنین نوع ابزار (نظیر پرسشنامه خودگزارش‌دهی یا مصاحبه) ذکر شود.
A2	ویژگی مورد سنجش	کدام شاخص یا ویژگی‌ها مورد مطالعه قرار گرفته است (به‌عنوان مثال، اگر تعداد زیادی از ویژگی‌ها بررسی شده‌اند).
A3	طراحی	نوع مطالعه مورد استفاده جهت سنجش ویژگی‌ها (نظیر طراحی آزمون بازآزمون، مطالعه طولی، هم‌گروهی، مقطعی یا تصادفی) و در صورت لزوم، سایر جزئیات طراحی مطالعه (مداخله / مواجهه، توصیف ابزارهای مقایسه، پیامدهایی غیر از سنجه‌های پیامدی گزارش‌شده توسط بیمار) ذکر شود.
A4	نمونه	معیارهای ورود / خروج ذکر شود. کلیات خصوصیات نمونه مورد مطالعه (نظیر شرایط، موقعیت جغرافیایی، زبان، سایر مشخصات پایه و جمعیت‌شناختی مرتبط) توصیف شود.
A5	مواد و روش‌ها	روش‌های بررسی هریک از ویژگی‌های مورد سنجش، از جمله آنالیزهای آماری، به‌طور مختصر توصیف شود.
A6	نتایج	نتایج اصلی کلیه ویژگی‌های موردسنجش همراه با شاخص‌های آماری برای هر نتیجه همراه با دقت اندازه‌گیری ذکر شود.
A7	بحث / نتیجه‌گیری	توصیف نتایج در زمینه شواهد موجود، نقاط قوت، محدودیت‌ها و نیاز به تحقیقات آینده پیرامون سنجه‌های پیامدی گزارش‌شده توسط بیمار مورد بررسی، به‌طور مختصر ذکر شود.
I1	نام و توصیف سنجه‌های پیامدی گزارش‌شده توسط بیمار مورد استفاده	نام اختصاصی، نوع، زبان و نسخه سنجه‌های پیامدی گزارش‌شده توسط بیمار مورد بررسی و چگونگی توسعه آن بیان شود. اهداف سازه مورد سنجش سنجه‌های پیامدی گزارش‌شده توسط بیمار و ابعاد مختلف آن توصیف شود؛ همچنین ساختار سنجه‌های پیامدی گزارش‌شده توسط بیمار شرح داده شود (نظیر تعداد عوامل، تعداد گویه‌ها و الگوریتم نمره‌دهی). دستورالعمل‌های مرتبط با ابزار (نظیر دوره‌های زمانی) و تعداد یا انواع دسته‌بندی‌های پاسخ‌دهی به ابزار ذکر شود. بیان شود که آیا سنجه‌های پیامدی گزارش‌شده توسط بیمار براساس مدل بازتابی (Reflective) یا تکوینی (Formative) طراحی شده است. توجه: این اطلاعات، ممکن است همچنین در بخش مواد و روش‌ها با جزئیات بیشتر ذکر شود.
I2	جمعیت هدف	جمعیت هدفی که سنجه‌های پیامدی گزارش‌شده توسط بیمار مختص آن گروه طراحی شده است، توصیف شود. نویسندگان باید ویژگی‌های مناسب و ضروری این جمعیت را بیان کنند.
I3	ارجاع به نسخه اصلی توسعه سنجه‌های پیامدی گزارش‌شده توسط بیمار	استاد به مقاله(های) اصلی توسعه سنجه‌های پیامدی گزارش‌شده توسط بیمار باید ارجاع شود و سایر استنادهای مرتبط با کیفیت خاص سنجه‌های پیامدی گزارش‌شده توسط بیمار تحت بررسی قرار گیرد.
I4	بیان دانش موجود و منطق	وضعیت علمی موجود در مورد ویژگی مورد اندازه‌گیری بیان شود (آنچه تاکنون شناخته شده است). نویسندگان باید یک مرور در ارتباط با شواهد موجود از نسخه خاص ابزار (نظیر نسخه کوتاه یا زبانی) فراهم کنند یا علت نیاز و اهمیت مطالعه جدید را توضیح دهند. منطق پیشنهاد مطالعه حاضر باید بیان شود.
I5	تعاریف	اصطلاحات تخصصی باید تعریف یا توضیح داده شوند.
I6	اهداف و فرضیه‌ها	هدف(های) اختصاصی پژوهش و فرضیه(های) مرتبط با سنجه‌های پیامدی گزارش‌شده توسط بیمار مورد بررسی، بیان شود.

توصیه‌های عمومی گزارش‌دهی مرتبط با کلیه مطالعات روان‌سنجی ابزار		
شماره گویه	نام گویه	توصیف گویه
GM1	طراحی مطالعه	عناصر کلیدی طرح مطالعه بیان شود.
GM2	شرکت‌کنندگان	نحوه انتخاب شرکت‌کنندگان توضیح داده شود؛ معیارهای ورود و خروج مطالعه ذکر شود (به‌عنوان مثال، معیارهای انتخاب شرکت‌کنندگان باید بازتابی از شرایط خاص سنجه‌های پیامدی گزارش‌دهی شده توسط بیمار باشد).
GM3	استفاده از سنجه‌های پیامدی گزارش‌دهی شده توسط بیمار	صراحتاً در مورد زمان و چگونگی استفاده از سنجه‌های پیامدی گزارش‌دهی شده توسط بیمار (نظیر محیط پژوهش) و روش جمع‌آوری داده‌ها (نظیر قلم-کاغذی یا الکترونیکی) توضیح داده شود.
GM4	روش‌های جمع‌آوری داده‌ها	سایر اطلاعات لازم پیرامون جمع‌آوری داده‌ها، روش‌های مواجهه (نظیر تخصیص نمونه‌ها جهت مداخله) و نقاط زمانی / پیگیری نمونه‌ها ارائه شود.
GM5	محاسبه توان / حجم نمونه	محاسبات توان مطالعه برای کلیه آنالیزهای مرتبط با ویژگی‌های اندازه‌گیری ارائه شود. اگر از قاعده سرانگشتی استفاده شده است، نحوه محاسبه و منبع استناد آن بیان شود.
GM6	تجزیه و تحلیل‌های آماری	تجزیه و تحلیل‌های آماری و آزمون‌های مرتبط با کلیه فرضیه‌ها و اهداف ویژگی‌های مورد اندازه‌گیری باید گزارش شود. همچنین نقطه برش اهمیت آماری باید ذکر شود (نظیر p-value کمتر از ۰/۰۵). کلیه تجزیه و تحلیل‌های مورد استفاده جهت تعیین بزرگی و جهت اثر نیز باید همراه با معیارهای دقت یا تغییرپذیری گزارش شود. به بسته آماری مورد استفاده در مطالعه اشاره شود.
GM7	مقادیر گمشده (Missing Data)	رویکرد یا برنامه جهت برخورد با مقادیر گمشده بیان شود.
GM8	آزمون‌های تعقیبی (Post-hoc Analysis)	تجزیه و تحلیل‌های اختصاصی که پس از دوره جمع‌آوری داده‌ها مورد استفاده قرار گرفته است (به‌عنوان مثال، اگر آنالیزها از نوع آزمون‌های تعقیبی بودند) همراه با دلایل منطقی انجام آزمون‌های تعقیبی باید توصیف شوند.
GR1	مقادیر گمشده	مقدار و علت از دست رفتن داده‌ها برای کلیه تجزیه و تحلیل‌های سنجه‌های پیامدی گزارش شده توسط بیمار (یا سایر پیامدهای اندازه‌گیری شده) و گروه‌های مرتبط باید توضیح داده شود.
GR2	مشخصات بیماران / شرکت‌کنندگان	مشخصات بیماران مورد مطالعه از جمله نمرات پایه سنجه‌های پیامدی گزارش شده توسط بیمار باید بیان شود.
GR3	حجم نمونه	اگر یک مطالعه شامل تجزیه و تحلیل‌هایی با حجم نمونه‌های متفاوت بودند، باید برای هر یک از آنالیزها، حجم نمونه گزارش شود.
D1	شواهد ویژگی‌های موردسنجش	برای هر یک از ویژگی‌های موردسنجش، نتایج باید با معیارهای اندازه‌گیری مطلوب (نظیر معیارهای کاسمین) مقایسه شود؛ همچنین کفایت ویژگی اندازه‌گیری شده مشخص شود. نکته: این اطلاعات ممکن است در بخش نتایج نیز با جزئیات بیشتری بیان شود (به‌عنوان مثال، در جدول).
D2	کاربردهای مرتبط	باید در ارتباط با کاربرد نتایج بحث شود.
D3	نقاط قوت و محدودیت‌ها	نقاط قوت و محدودیت‌های مطالعه باید مورد بحث قرار گیرند. به‌عنوان مثال، اگر سوگیری‌های بالقوه قابل توجهی در مطالعه وجود دارد که می‌تواند بر نتایج تأثیر گذارد، درباره‌شان بحث شود.
D4	تعمیم‌پذیری	مسائل تعمیم‌پذیری مرتبط با نتایج سنجه‌های پیامدی گزارش شده توسط بیمار باید بحث شوند. به‌عنوان مثال، اگر نتایج می‌تواند به سایر جمعیت‌ها تعمیم داده شود، ذکر شود.
D5	تغییرات ابزار	در مورد نیاز به تغییر در ابزار سنجه‌های پیامدی گزارش شده توسط بیمار موجود یا توسعه نسخه جدید، بحث شود. اگر نویسندگان به این نتیجه رسیدند که یکی از ویژگی‌های اندازه‌گیری کافی نیست، تغییرات پیشنهاد داده شود. یا اگر واقعاً ضعیف است، می‌توان توقف استفاده از سنجه‌های پیامدی گزارش شده توسط بیمار (به‌طور کلی یا در یک جمعیت خاص) پیشنهاد داده شود.
D6	تحقیقات آینده	نوع تحقیقات خاص مورد نیاز برای پاسخ به سؤالات جدید ناشی از نتایج مطالعه حاضر برای بررسی ویژگی‌های اندازه‌گیری سنجه‌های پیامدی گزارش شده توسط بیمار گزارش شود.
C1	نتیجه‌گیری	نتیجه‌گیری کلی برای هر ویژگی موردسنجش و استفاده از سنجه‌های پیامدی گزارش شده توسط بیمار بیان شود.

توصیه‌های عمومی گزارش‌دهی مرتبط با کلیه مطالعات روان‌سنجی ابزار			
شماره گویه	نام گویه	توصیف گویه	
O1	تضاد منافع	هرگونه تضاد منافع مرتبط با سنجه‌های پیامدی گزارش‌شده توسط بیمار، ذکر شود (به‌عنوان مثال، همکاری یکی از طراحان سنجه‌های پیامدی گزارش‌شده توسط بیمار، به‌عنوان نویسنده یا سازمان تأمین‌کننده)	بخش گزارش: سایر اطلاعات
CV1	مرتبط بودن	گزارش شود که آیا از بیماران و یا متخصصان، پیرامون ارتباط هر گویه با تجارب یا موقعیت آن‌ها سؤال شده است؟ چگونه؟	
CV2	جامعیت	گزارش شود که آیا از بیماران و یا متخصصان، پیرامون گنجاندن شدن کلیه مفاهیم کلیدی سازه نظرخواهی شده است؟ چگونه؟	
CV3	ادراک‌پذیری	گزارش شود که آیا درک دستورالعمل‌های سنجه‌های پیامدی گزارش‌شده توسط بیمار، گویه‌ها، گزینه‌های پاسخ‌دهی و دوره یادآوری ارزیابی شده است؟ چگونه؟	
CV4	نتایج مرتبط	اگر موارد خاصی در ارتباط با جامعیت و یا زمینه مورد استفاده توسط بیماران و یا متخصصان در نظر گرفته شده است، گزارش شود.	توصیه‌های اختصاصی گزارش‌دهی مطالعات مربوط به روایی محتوا
CV5	گزینه‌های پاسخ‌دهی و دوره یادآوری	گزارش شود که آیا گزینه‌های پاسخ‌دهی و دوره یادآوری متناسب با بیماران و یا متخصصان در نظر گرفته شده است؟	
CV6	نتایج جامعیت	گزارش شود که آیا بیماران و یا متخصصان، کلیه مفاهیم کلیدی را در سنجه‌های پیامدی گزارش‌شده توسط بیمار، لحاظ کرده‌اند؟	
CV7	نتایج ادراک‌پذیری (Comprehensibility)	گزارش شود که آیا بیماران، دستورالعمل‌ها، گویه‌ها و گزینه‌های پاسخ‌دهی را درک کرده‌اند؟	
SV1	تحلیل عاملی: نظریه آزمون کلاسیک (CTT)	جزئیات روش‌ها و نتایج آنالیز عاملی اکتشافی یا تأییدی گزارش شود. منطق انجام آنالیز عاملی اکتشافی بیان شود (به‌عنوان مثال، عدم وجود فرضیه‌های پیش‌بین مشخص). برای آنالیز عاملی تأییدی، ساختار عاملی مدل‌های آزمون‌شده توصیف و توجیه شود. روش‌ها و نتایج بررسی پیش‌فرض‌ها، روش تخمین، شاخص‌های برازش و نقاط برش برای برازش مدل شامل بارهای عاملی بهترین مدل برازش‌شده، باید توصیف شوند.	توصیه‌های اختصاصی گزارش‌دهی مطالعات مربوط به روایی ساختاری
SV2	تجزیه و تحلیل‌های نظریه آیتم-پاسخ (IRT)	نوع مدل سؤال-پاسخ یا راش باید گزارش شود. همچنین روش تخمین، روش‌ها و نتایج بررسی پیش‌فرض‌ها (تک‌بعدی (به تحلیل عاملی مراجعه کنید)، وابستگی (نظیر همبستگی‌های باقی‌مانده)، یکنواختی (نظیر مقیاس مومن (Mokken Scaling))، شاخص‌های برازش و نقاط برش برای برازش مدل و سایر پارامترهای گویه‌ها) باید گزارش شوند.	
IC1	واحد اندازه‌گیری	روش‌ها و نتایج سنجش همسانی درونی برای هر مقیاس تک‌بعدی یا زیرمقیاس‌ها گزارش شود. کلیه شواهد یا پیش‌فرض‌های مرتبط با تک‌بعدی بودن ابزار ذکر شود.	
IC2	نمرات پیوسته	شاخص‌های آلفای کرونباخ یا امگا گزارش شود. سایر شاخص‌های محاسبه‌شده برای سنجش همسانی درونی نمرات پیوسته را ذکر شود.	توصیه‌های اختصاصی گزارش‌دهی مطالعات مربوط به همسانی درونی
IC3	نمرات دوتایی	شاخص‌های آلفای کرونباخ یا ضریب کودر-ریچاردسون (Kuder-Richardson Coef-ficient) گزارش شود. سایر شاخص‌های محاسبه‌شده برای سنجش همسانی درونی نمرات دوتایی ذکر شود.	
CCV1	گروه(های) مقایسه	ویژگی‌های گروه(های) مقایسه، از جمله حجم نمونه در هر گروه، گزارش شود.	
CCV2	تحلیل عاملی: نظریه آزمون کلاسیک (CTT)	جزئیات روش‌ها و نتایج تحلیل عاملی تأییدی گروه‌های چندگانه، آنالیزهای رگرسیون لجستیک (Logistic Regression) یا سایر آنالیزهای انجام‌شده، گزارش شود. کلیات مدل‌های آزمون‌شده از جمله بارهای عاملی، توابع و واریانس‌های تحلیل عاملی تأییدی توصیف و توجیه شود. روش‌ها و نتایج ارزیابی پیش‌فرض‌ها باید توصیف شود. معیارهای تغییرناپذیری باید تعریف شوند. روش تخمین، شاخص‌های برازش و معیارهای سنجش تغییرناپذیری ابزار بیان شود.	توصیه‌های اختصاصی گزارش‌دهی مطالعات مربوط به روایی بین فرهنگی / ثبات اندازه‌گیری
CCV3	تجزیه و تحلیل‌های نظریه آیتم-پاسخ (IRT)	نوع مدل سؤال-پاسخ یا راش باید گزارش شود. همچنین روش تخمین، روش‌ها و نتایج بررسی پیش‌فرض‌ها (تک‌بعدی (به تحلیل عاملی مراجعه کنید)، وابستگی (نظیر همبستگی‌های باقی‌مانده)، یکنواختی (نظیر مقیاس مومن)) باید گزارش شوند. بسته‌های آماری مورد استفاده، روش تخمین، معیارهای استفاده‌شده جهت شناسایی کنش افتراقی (Differential Item Functioning; DIF) سوالات و روش‌ها و نتایج کلیه مقایسه‌های مدل باید توصیف شوند.	

توصیه‌های عمومی گزارش‌دهی مرتبط با کلیه مطالعات روان‌سنجی ابزار

شماره گویه	نام گویه	توصیف گویه
R1	استفاده از سنجه‌های پیامدی گزارش‌شده توسط بیمار	تعداد کل اندازه‌گیری‌های انجام‌شده گزارش شود. ذکر شود که آیا اندازه‌گیری‌ها با استفاده از همان سنجه‌های پیامدی گزارش‌شده توسط بیمار، روی نمونه‌های قبلی انجام شده است. همچنین فرایند انجام اندازه‌گیری‌ها در بیماران باید توضیح داده شود، از جمله اینکه اندازه‌گیری توسط چه کسی (آیا بیمار آن را تکمیل کرده است یا یک واسطه یا نماینده وجود داشته است)، چه زمانی، چگونه و با چه فاصله زمانی انجام شده است. این موارد باید شامل فاصله زمانی بین اندازه‌گیری‌های مکرر (به‌عنوان مثال، آیا بیمار ثابت بوده است)، نوع آزمون (نظیر پرسش‌نامه خودتکمیل‌کننده یا ابزار مصاحبه‌محور)، محیط اندازه‌گیری (نظیر بیمارستان یا منزل) باشد و دستورالعمل‌های تکمیل ابزار باید در گزارش ذکر شود. در صورت لزوم، سایر ابزارها یا اندازه‌گیری‌های همراه با سنجش سنجه‌های پیامدی گزارش‌شده توسط بیمار و مستقل بودن دفعات تکمیل ابزارها از یکدیگر (آیا سنجه‌های پیامدی گزارش‌شده توسط بیمار، بدون اطلاع از نمرات قبلی در نمونه‌ها تکمیل شده است) ذکر شود.
R2	آنالیزهای آماری	کلیه تجزیه‌وتحلیل‌های آماری و نتایج ارزیابی‌های (های) پایایی باید بیان شود و کاربرد آن‌ها مورد قضاوت قرار گیرد (نظیر مدل ضریب همبستگی درون‌خوشه‌ای یا نوع ضریب کاپای استفاده‌شده). همچنین مؤلفه‌های واریانس و وزن‌دهی نمرات ترتیبی (نظیر وزن‌دهی خطی یا چهاربر (Quadratic) بیان شود.
R3	روش‌های بهبود پایایی	روش‌های بهبود پایایی در مطالعه نظیر محدودسازی نمونه، آموزش محققان، استانداردسازی روش‌ها و استفاده از میانگین اندازه‌گیری‌های مکرر گزارش شود.
ME1	استفاده از سنجه‌های پیامدی گزارش‌شده توسط بیمار	تعداد کل اندازه‌گیری‌های انجام‌شده گزارش شود. ذکر شود که آیا اندازه‌گیری‌ها با استفاده از همان سنجه‌های پیامدی گزارش‌شده توسط بیمار روی نمونه‌های پیشین انجام شده است. همچنین فرایند انجام اندازه‌گیری‌ها در بیماران باید توضیح داده شود، از جمله اینکه اندازه‌گیری توسط چه کسی (آیا بیمار آن را تکمیل کرده است یا یک واسطه یا نماینده وجود داشته است)، چه زمانی، چگونه و با چه فاصله زمانی انجام شده است؛ این موارد باید شامل فاصله زمانی بین اندازه‌گیری‌های مکرر (به‌عنوان مثال، آیا بیمار ثابت بوده است)، نوع آزمون (نظیر پرسش‌نامه خودتکمیل‌کننده یا ابزار مصاحبه‌محور)، محیط اندازه‌گیری (نظیر بیمارستان یا منزل) و دستورالعمل‌های تکمیل ابزار باشد. در صورت لزوم، سایر ابزارها یا اندازه‌گیری‌های همراه با سنجش سنجه‌های پیامدی گزارش‌شده توسط بیمار و استقلال تکمیل ابزارها (آیا سنجه‌های پیامدی گزارش‌شده توسط بیمار بدون اطلاع از نمرات قبلی در نمونه‌ها تکمیل شده است) ذکر شود.
ME2	آنالیزهای آماری	کلیه تجزیه‌وتحلیل‌های آماری و نتایج ارزیابی‌های (های) خطای اندازه‌گیری باید بیان شود و کاربرد آن‌ها به‌ویژه برای نمرات پیوسته، خطای استاندارد اندازه‌گیری (مدل محاسبه SEM (نظیر توافق SEM یا پایایی SEM)، حداقل تغییر قابل تشخیص (فرمول اختصاصی مورد استفاده، شامل مدل SEM) یا محدودیت‌های توافق مورد قضاوت قرار گیرد. همچنین باید برای نمرات دوتایی / اسمی / ترتیبی، حاشیه‌ها (داده‌های خام) و درصد توافق (نظیر مثبت و منفی) گزارش شود.
CriV1	معیار	جزئیات معیار مورد استفاده و علت استفاده از آن گزارش شود. توجیه استاندارد طلایی باید بیان شود. همچنین، در صورت لزوم، توصیف شود که چرا و چگونه روایی ملاکی انجام شده است (به‌عنوان مثال، اگر مستقل از سنجه‌های پیامدی گزارش‌شده توسط بیمار است).
CriV2	نمرات پیوسته	همبستگی‌ها (زمانی که معیار دارای امتیازات پیوسته است) یا سطح زیر منحنی راک (Receiver Operating Characteristic) (زمانی که معیار دوتایی است) گزارش شود.
CriV3	نمرات طبقه‌بندی‌شده	توصیف شود که چرا و چگونه نمرات سنجه‌های پیامدی گزارش‌شده توسط بیمار به ۲ یا چند دسته تقسیم شده است. شاخص‌های حساسیت و ویژگی گزارش شود.

توصیه‌های اختصاصی  
گزارش‌دهی مطالعات مربوط  
به پایایی

توصیه‌های اختصاصی  
گزارش‌دهی مطالعات مربوط  
به خطای اندازه‌گیری

توصیه‌های اختصاصی  
گزارش‌دهی مطالعات مربوط  
به روایی ملاکی

توصیه‌های عمومی گزارش‌دهی مرتبط با کلیه مطالعات روان‌سنجی ابزار		
شماره گویه	نام گویه	توصیف گویه
ConV1	ابزار(های) مقایسه‌کننده	ابزارهای مقایسه‌کننده باید از نظر ساختار(هایی) که قصد اندازه‌گیری آن‌ها را دارند، به‌طور مناسب توصیف شوند. ویژگی‌های اندازه‌گیری ابزارهای مقایسه‌کننده و معیارها یا داده‌های مرتبط گزارش شوند.
ConV2	گروه(های) مقایسه‌کننده	ویژگی‌های گروه‌های مورد مقایسه شامل حجم نمونه در هر گروه، گزارش شود.
ConV3	فرضیه‌ها	کلیه فرضیه‌ها، از جمله جهت و بزرگی همبستگی‌های مورد انتظار بین سنجه‌های پیامدی گزارش شده توسط بیمار و سایر ابزارهای اندازه‌گیری، یا جهت و بزرگی تفاوت‌ها در نمرات سنجه‌های پیامدی گزارش شده توسط بیمار، بین گروه‌ها گزارش شود.
ConV4	تجزیه و تحلیل‌های آماری	کلیه روش‌های تجزیه و تحلیل و نتایج مورد استفاده برای آزمون هر فرضیه، گزارش شود.
ConV5	نتایج	گزارش شود که کدام نتایج خاص با فرضیه آن مطابقت دارد.
Resp1	ابزار(های) مقایسه‌کننده	ابزارهای مقایسه‌کننده باید از نظر ساختار(هایی) که قصد اندازه‌گیری آن‌ها را دارند، به‌طور مناسب توصیف شوند. ویژگی‌های اندازه‌گیری ابزارهای مقایسه‌کننده و استانداردها یا داده‌های مرتبط گزارش شود.
Resp2	گروه(های) مقایسه‌کننده	ویژگی‌های گروه‌های مورد مقایسه شامل حجم نمونه در هر گروه، گزارش شود.
Resp3	فرضیه‌ها	کلیه فرضیه‌ها، از جمله جهت و بزرگی همبستگی‌های مورد انتظار بین تغییرات در سنجه‌های پیامدی گزارش شده توسط بیمار و تغییر در سایر ابزارهای اندازه‌گیری، یا جهت و بزرگی تفاوت‌ها در تغییرات نمره سنجه‌های پیامدی گزارش شده توسط بیمار بین گروه‌ها گزارش شود.
Resp4	روش‌های اندازه‌گیری	گزارش شود که آیا اندازه‌گیری‌ها با استفاده از ابزارهای مشابه روی نمونه‌های یکسان انجام شده است. روش‌های اندازه‌گیری از جمله فواصل زمانی بین ابزارهای اندازه‌گیری، بیان شود.
Resp5	طول مدت فاصله	طول مدت فاصله بین نقاط زمانی مطالعه باید توضیح داده شوند.
Resp6	مداخله / مواجهه	در صورت لزوم، مداخله انجام شده یا مواجهه در دوره فاصله توصیف شود.
Resp7	بیماران تغییر یافته	نسبت بیماران که بهبود یافته یا بدتر شده‌اند (و جزئیات هر مرجع (Anchor) مورد استفاده) در ساختار اندازه‌گیری شده کلیه سنجه‌های پیامدی گزارش شده توسط بیمار گزارش شود. هرگونه تغییر در نمرات سنجه‌های پیامدی گزارش شده توسط بیمار، در جمعیت هدف برای کاربرد پژوهش مربوط به فرضیه‌های از پیش تعریف شده، گزارش شود.
Resp8	تجزیه و تحلیل‌های آماری	کلیه روش‌های آماری و نتایج مورد استفاده برای آزمون هر فرضیه، گزارش شود.
Resp9	نتایج	گزارش شود که کدام نتایج خاص با فرضیه آن مطابقت دارد.

نظیر سنجه‌های پیامدی گزارش شده توسط پزشک<sup>۱۲</sup>، سنجه‌های پیامدی عملکرد محور<sup>۱۳</sup>، سنجه‌های پیامدی گزارش شده توسط ناظر<sup>۱۴</sup> و تست‌های آزمایشگاهی، منابع تأثیرگذار و تغییردهنده نمرات بیشتری دارند، بنابراین معمولاً از ساختار پیچیده‌تری در روان‌سنجی بهره‌مند هستند [۲۴]. در این زمینه، مطالعاتی به‌عنوان مطالعات باکیفیت مطرح هستند که بازتابی از میزان خطای اندازه‌گیری و قابلیت اطمینان ابزار در شرایط مختلف را فراهم کنند؛ بنابراین ارزیابی کیفیت مطالعات در زمینه قابلیت اطمینان و خطای اندازه‌گیری همواره به‌عنوان یکی از چالش‌های پژوهشگران حوزه روان‌سنجی ابزار مطرح است [۲۴، ۲۰]. در این راستا، پژوهشگران کاسمین در سال ۲۰۲۰ با هدف تعیین پایایی

12. Clinician-Reported Outcome Measures (ClinROMs)
13. Performance-Based Outcome Measures (PerFOMs)
14. Observer-Reported Outcome Measures (ObsROMs)

جهت ارزیابی کیفیت روش‌شناسی هر مطالعه با استفاده از چک‌لیست خطر سوگیری کاسمین، هر استاندارد ارائه شده در طیف بسیار خوب، کافی، مشکوک و ناکافی ارزیابی می‌شود. در نهایت پایین‌ترین رتبه‌بندی از هر استاندارد ارائه شده در هر کادر جهت ارزیابی کلی در نظر گرفته می‌شود. در مواردی که استاندارد مورد نظر برای مطالعه قابل استفاده نیست، از گزینه «غیر قابل اجرا» استفاده می‌شود [۱، ۱۸، ۲۳].

### ابزار خطر سوگیری کاسمین جهت ارزیابی کیفیت مطالعات با تأکید بر پایایی یا خطای اندازه‌گیری

یکی از مهم‌ترین ویژگی‌های سنجه‌های پیامدی گزارش شده توسط بیمار، میزان تغییرات واقعی است که با تکرار سنجش توسط ابزار رخ می‌دهد. از آنجایی که گروهی از ابزارهای سنجش سلامت

یا خطای اندازه‌گیری ابزارهای پیچیده، نسخه دیگری از ابزار خطر سوگیری را براساس روش دلفی ارائه دادند.

این ابزار از ۲ بخش تشکیل شده است. هدف از بخش اول چک‌لیست، کمک به درک و آگاهی بخشی به پژوهشگران در ارتباط با پتانسیل ابزار در سنجش ویژگی‌های پایایی و خطای اندازه‌گیری است. برای درک این موضوع باید ۷ عنصر پژوهش، مشخص و واضح شود. در این راستا عناصر تشکیل‌دهنده سؤال پژوهش باید تعیین شود که شامل [۱] نام ابزار سنجش پیامد، [۲] نسخه ابزار یا روش عملیاتی‌سازی پروتکل اندازه‌گیری، [۳] ساختار مورد اندازه‌گیری توسط ابزار، [۴] تشخیص ویژگی مرتبط با پایایی (شاخص‌های نسبی نظیر ضریب همبستگی درون‌خوشه‌ای<sup>۱۵</sup>، ضریب تعمیم‌پذیری فی  $\Phi$  یا کاپا  $K$ ) یا خطای اندازه‌گیری (شاخص‌های مطلق نظیر خطای استاندارد اندازه‌گیری<sup>۱۶</sup>، محدودیت‌های توافق<sup>۱۷</sup>، حداقل تغییر قابل تشخیص<sup>۱۸</sup> یا درصد توافق خاص) در ابزار، [۵] تعیین اجزایی از ابزار اندازه‌گیری که تکرار خواهند شد (به‌ویژه زمانی که فقط بخشی از ابزار تکرار می‌شود)، [۶] تعیین منابعی که تغییر خواهند کرد (نظیر زمان یا رخداد، متخصصان، وسایل و یا سایر اجزای اندازه‌گیری) و [۷] مشخصات جمعیت مورد مطالعه، هستند [۲۰].

بخش دوم چک‌لیست جهت ارزیابی کیفیت یک مطالعه پیرامون ویژگی‌های پایایی یا خطای اندازه‌گیری طراحی شده است. این بخش شامل استانداردهایی است که قابلیت اعتماد به ابزار را بررسی می‌کند. سنجش پایایی با استفاده از ۶ استاندارد در مورد الزامات طراحی و ۳ استاندارد در مورد روش‌های آماری ترجیحی انجام می‌شود. سنجش خطای اندازه‌گیری با استفاده از همان ۶ استاندارد پیرامون الزامات طراحی و ۲ استاندارد دیگر در مورد روش‌های آماری ترجیحی (متفاوت از استانداردهای سنجش پایایی) صورت می‌گیرد؛ چراکه طراحی مطالعه برای ویژگی‌های پایایی و خطای اندازه‌گیری مشابه است و از داده‌های یکسانی در این مرحله استفاده می‌شود، اما روش‌های آماری مورد استفاده متفاوت خواهد بود. در این بخش، همانند چک‌لیست خطر سوگیری کاسمین، هر استاندارد در طیف بسیار خوب، کافی، مشکوک، ناکافی و غیرقابل اجرا ارزیابی می‌شود [۲۰].

### چک‌لیست طراحی مطالعه کاسمین

چک‌لیست طراحی مطالعه کاسمین [۲۱] براساس نسخه اصلی چک‌لیست کاسمین [۸] جهت استفاده پژوهشگران حین طراحی مطالعات با تأکید بر ویژگی‌های روان‌سنجی سنجش‌های پیامدی گزارش شده توسط بیمار در سال ۲۰۱۹ معرفی شده است [۲۱].

15. Intraclass Correlation Coefficient (ICC)
16. Standard Error of Measurement (SEM)
17. Limits of Agreement (LoA)
18. Smallest Detectable Change (SDC)

این چک‌لیست از ۱۰ بخش تشکیل شده است. در هر بخش از این چک‌لیست، استانداردهایی برای مطالعات خاص با تأکید بر ویژگی‌های روان‌سنجی ابزار مطرح می‌شود که شامل [۱] توصیه‌های کلی برای طراحی یک مطالعه سنجش ویژگی‌های روان‌سنجی ابزار، [۲] روایی محتوا، [۳] روایی ساختاری، [۴] همسانی درونی، [۵] روایی بین فرهنگی / ثبات اندازه‌گیری، [۶] پایایی و خطای اندازه‌گیری، [۷] روایی ملاکی (باتوجه به ماهیت خود گزارش‌دهی ابزارهای سنجش‌های پیامدی گزارش شده توسط بیمار، هیچ استاندارد طلایی برای این ابزارها وجود ندارد و صرفاً می‌توان از نسخه طولانی برای بررسی ویژگی‌های روان‌سنجی نسخه کوتاه همان ابزار استفاده کرد)، [۸] آزمون فرضیه جهت سنجش روایی سازه (شامل روایی هم‌گرا و روایی واگرا یا مقایسه گروه‌های شناخته‌شده)، [۹] و قابلیت پاسخ‌گویی (رویکرد معیاری شامل مقایسه با استاندارد طلایی و رویکرد سازه‌ای شامل مقایسه با سایر ابزارهای سنجش پیامد، مقایسه بین گروه‌های شناخته‌شده و مداخلات قبل و بعد جهت آزمون فرضیه) و [۱۰] فرایند ترجمه است. با وجود اینکه فرایند ترجمه، به‌عنوان یک ویژگی اندازه‌گیری مطرح نیست، اما یک فرایند ترجمه خوب، به ارائه نسخه معتبرتر از سنجش‌های پیامدی گزارش شده توسط بیمار منجر می‌شود. بنابراین استانداردهای ترجمه یک سنجش پیامدی در فرایند ترجمه ابزار نیز در قالب بخش دهم چک‌لیست ارائه شده است [۱۸، ۲۱]. در بخش‌های ۳ و ۴ (روایی ساختاری و همسانی درونی)، ابزارهایی مورد بررسی قرار می‌گیرند که براساس مدل بازتابی<sup>۱۹</sup> طراحی شده باشند و از نوع تک‌وجهی<sup>۲۰</sup> نباشند. همچنین در صورت چندبعدی بودن، استانداردهای مذکور باید برای هر یک از خرده مقیاس‌ها نیز بررسی شوند [۱۸، ۲۱].

در این چک‌لیست، مطابق مقیاس رتبه‌بندی چک‌لیست خطر سوگیری کاسمین، هر استاندارد در طیف بسیار خوب، کافی، مشکوک، ناکافی و غیرقابل اجرا رتبه‌بندی می‌شود؛ این رتبه‌بندی با هدف درک بهتر کیفیت طراحی مطالعه صورت می‌گیرد تا مسائل مهم و جزئیات ضروری در طراحی مطالعات روان‌سنجی فراموش نشوند [۲۱].

### راهنمای گزارش کاسمین

پژوهشگران کاسمین با همکاری دانشگاه میشیگان در سال ۲۰۲۱، دستورالعملی را برای گزارش مطالعات مرتبط با سنجش ویژگی‌های روان‌سنجی ابزارهای پیامدی گزارش شده توسط بیمار فراهم کرده‌اند که تحت عنوان راهنمای گزارش کاسمین منتشر شده است [۲۵]. استفاده از این چک‌لیست در گزارش مطالعات روان‌سنجی حائز اهمیت است؛ چراکه با بهبود گزارش‌دهی، شفافیت ویژگی‌های ابزارها افزایش می‌یابد و می‌توان با اعتماد بیشتری از

19. Reflective
20. Formative

یافته‌های مطالعات و ابزار مورد نظر استفاده کرد [۲۵، ۲۱].

کاسمین توسعه یافته است و تعداد زیادی از توسعه‌دهندگان سنجه‌های پیامدی گزارش شده توسط بیمار مورد بررسی در مطالعه شرکت نداشتند و یا انصراف دادند [۲۷]. این چک‌لیست به‌عنوان یک ابزار مبتنی بر نظرات در نظر گرفته می‌شود، نه مبتنی بر شواهد؛ بنابراین این انتقاد به چک‌لیست وارد است که ممکن است برخی از روش‌ها یا ویژگی‌ها در نظر گرفته نشده باشد. امید است که با به‌روزرسانی نسخه‌های جدیدتر و کاربست روش‌های جامع‌تر در توسعه چک‌لیست، این مورد برطرف شود.

از سوی دیگر، در نظر گرفتن همیشگی چک‌لیست کاسمین به‌عنوان یک معیار جامع در ارزیابی کیفیت مطالعات روان‌سنجی، ممکن است به توقف یا کند شدن روند توسعه علم اندازه‌گیری منجر شود. چراکه طراحان ابزار، همچنان در جست‌وجوی اجرای استانداردهای گذشته در مطالعات هستند و برخی از انواع سنجه‌های پیامدی و ویژگی‌های روان‌سنجی مرتبط با آن‌ها در نظر گرفته نخواهد شد [۲۷].

همانند سایر ابزارها، پایایی چک‌لیست‌های کاسمین نیز باید توسط پژوهشگران بررسی شود. با توجه به انجام عمده مطالعات سنجش پایایی این چک‌لیست توسط محققان اصلی، به نظر می‌رسد بهتر است این مهم توسط پژوهشگران بی‌طرف انجام شود [۲۷]. در این زمینه، اسکرز ضمن بیان نکات مثبت در مورد چک‌لیست کاسمین، ادعا می‌کند که قابلیت اطمینان بین ارزیابان در نسخه اولیه چک‌لیست کاسمین پایین است و درصد توافق رتبه‌بندی‌ها برای حدود یک‌سوم توصیه‌ها کمتر از ۸۰ درصد گزارش شده است [۲۸]. بنابراین با توجه به معرفی نسخه‌های تجدیدنظر شده، لازم است تا مطالعات بیشتری در این زمینه انجام شود.

یکی دیگر از محدودیت‌های مطرح‌شده، تخصصی بودن چک‌لیست‌های کاسمین و به تبع آن، نیاز به زمان زیاد جهت آموزش نحوه تکمیل چک‌لیست‌ها و ارزیابی مطالعات است؛ بنابراین متخصصان حوزه‌هایی به‌جز روان‌سنجی نظیر پزشکان و پیراپزشکانی که بر بهبود کیفیت ارائه مراقبت‌ها متمرکز هستند، قادر به کاربست این چک‌لیست نخواهند بود و مسئولیت ارزیابی کیفیت مطالعات روان‌سنجی، صرفاً بر عهده محققان حوزه روان‌سنجی ابزار خواهد بود.

از نکات مهمی که باید در حوزه سنجش ویژگی‌های روان‌سنجی ابزار به آن توجه داشت، عدم یکسان بودن اهمیت ویژگی‌های روان‌سنجی در تصمیم‌گیری‌های بالینی است. بنابراین به نظر می‌رسد بهتر است برای تصمیم‌گیری‌ها یا اهداف بالینی معیارهای مناسب‌تری تعریف و ارائه شود [۲۶]. همچنین چک‌لیست کاسمین، برخی از شاخص‌های مهم کیفیت روان‌سنجی ابزار نظیر تولید و کاهش گویه و اعتبار داخلی را در نظر نگرفته یا به اندازه کافی توصیف نکرده است [۲۹]. همچنین با توجه به اهمیت

این راهنما، مجموعه‌ای از استانداردها پیرامون ویژگی‌های روان‌سنجی در مطالعات مرتبط ارائه می‌دهد که شامل توصیه‌های عمومی برای کلیه مطالعات روان‌سنجی و توصیه‌های اختصاصی در ارتباط با هر یک از ویژگی‌های روان‌سنجی می‌شود. توصیه‌های عمومی شامل استانداردهایی برای بهبود گزارش‌دهی در عنوان (۳ مورد)، چکیده (۷ مورد)، مقدمه (۶ مورد)، مواد و روش‌ها (۸ مورد)، نتایج (۳ مورد)، بحث (۶ مورد)، نتیجه‌گیری (۱ مورد) و سایر اطلاعات لازم (۱ مورد) در ارتباط با تضاد منافع است. توصیه‌های اختصاصی، استانداردهای گزارش‌دهی مربوط به روایی محتوا (۷ مورد)، روایی ساختاری (۲ مورد)، همسانی درونی (۳ مورد)، روایی بین فرهنگی / ثبات اندازه‌گیری (۳ مورد)، پایایی (۳ مورد)، خطای اندازه‌گیری (۲ مورد)، روایی ملاکی (۳ مورد)، آزمون فرضیه جهت روایی سازه (۵ مورد) و قابلیت پاسخ‌گویی (۹ مورد) را شامل می‌شوند.

باتوجه به اهمیت این موضوع، مقاله حاضر به ارائه نسخه فارسی این راهنما پرداخته است که در جدول شماره ۲ ارائه شده است [۲۵].

## بحث

مطالعه حاضر به بررسی و معرفی آخرین نسخه چک‌لیست‌های کاسمین پرداخته است. همگام با توسعه علم روان‌سنجی ابزار، پژوهشگران کاسمین نیز با هدف تقویت پایایی و سهولت استفاده توسط محققان به تغییر نسخه اصلی چک‌لیست پرداختند. این تغییرات به تفکیک چک‌لیست به ۴ نسخه متفاوت و بهبود ساختاری ابزار شد. با وجود تغییرات ایجادشده در ابزار، همچنان چالش‌هایی پیرامون چک‌لیست‌ها مطرح است.

ارزیابی‌های صورت‌گرفته توسط چک‌لیست کاسمین، صرفاً نشان‌دهنده کیفیت مطالعات روان‌سنجی ابزار است و نباید به‌عنوان معیاری جهت سنجش کیفیت یک ابزار اندازه‌گیری مورد استفاده قرار گیرد. به عبارت دیگر، در صورت عدم بررسی یا شواهد ناکافی از سنجش یک یا چند ویژگی روان‌سنجی خاص در یک مطالعه، کیفیت مطالعه پایین در نظر گرفته می‌شود؛ چراکه کلیه معیارهای چک‌لیست کاسمین را برآورده نمی‌کند [۲۶]. بنابراین ممکن است که تصمیم‌گیری‌های بالینی توسط متخصصان علوم بالینی نظیر پزشکان و پیراپزشکان را تحت تأثیر قرار دهد. به نظر می‌رسد بهتر است که این موضوع، در کلیه چک‌لیست‌ها برای محققانی که قصد استفاده از این چک‌لیست‌ها را دارند، شفاف‌سازی شود.

سنجه‌های پیامدی گزارش‌شده توسط بیمار شامل طیف وسیعی از ابزارها هستند که هر کدام ممکن است به روش‌شناسی متفاوتی جهت توسعه ابزار نیاز داشته باشند. از آنجایی که چک‌لیست کاسمین براساس تخصص و تجربه اعضای کمیته

### حامی مالی

مطالعه حاضر، فاقد حامی مالی بوده است.

### مشارکت نویسندگان

تمام نویسندگان در طراحی، اجرا و نگارش کلیه بخش‌های این مطالعه مشارکت داشته‌اند.

### تعارض منافع

نویسندگان هیچ تعارض منافی در این مطالعه نداشته‌اند.

### تشکر و قدردانی

بدین وسیله از کلیه محققان به‌ویژه پژوهشگران کاسمین که مقالات آن‌ها در این مطالعه مورد نقد و بررسی قرار گرفت و همچنین کمیته اخلاق دانشگاه علوم پزشکی کاشان تشکر و قدردانی می‌شود.

جزئیات مورد استفاده در بخش روایی ساختاری، در چک‌لیست کاسمین مسائل مهمی نظیر انتخاب روش استخراج عوامل در تحلیل عاملی اکتشافی (محوریابی مؤلفه‌های اصلی<sup>۲۱</sup> یا تحلیل عامل مشترک<sup>۲۲</sup>، انتخاب و توجیه روش چرخش (مثل چرخش متعامد<sup>۲۳</sup>) یا مایل<sup>۲۴</sup> و یا تصمیم‌گیری در مورد تعداد عوامل مرتبط، در نظر گرفته نشده است [۳۰].

باتوجه به جامعیت چک‌لیست‌های کاسمین، می‌توان در مواردی از این ابزارها جهت ارزیابی کیفیت طراحی، گزارش و نقد سنجه‌های پیامدی گزارش شده توسط غیر بیماران (ابزارهایی به‌جز سنجه‌های پیامدی گزارش شده توسط بیمار) نیز استفاده کرد [۱۳] که از نقاط قوت این چک‌لیست‌ها به شمار می‌رود.

### نتیجه‌گیری

باتوجه به نیاز به استفاده از سنجه‌های پیامدی معتبر و قابل اطمینان در تحقیقات و بالین و همچنین تعدد روزافزون ابزارها پیرامون هریک از سازه‌های مرتبط با سلامتی، نیاز به وجود یک مقیاس جهت ارائه استانداردهای مرتبط با سنجش ویژگی‌های روان‌سنجی ابزارها بیش از پیش احساس می‌شود که چک‌لیست‌های کاسمین، این مهم را با ارائه شاخص‌های منطقی تا حدود زیادی برطرف می‌کنند. با وجود مطرح شدن برخی از محدودیت‌ها پیرامون این چک‌لیست، چک‌لیست‌های کاسمین می‌توانند تا حدود زیادی در بررسی‌های سیستماتیک کیفیت مطالعات روان‌سنجی، طراحی مطالعات و یا گزارش مقالات روان‌سنجی ابزار توسط پژوهشگران، داوران و ویراستاران مجلات استفاده شوند.

باتوجه به اهمیت کیفیت طراحی، گزارش و نقد مطالعات روان‌سنجی ابزار، پیشنهاد می‌شود که محققان با شیوه ترجمه مستقیم و غیرمستقیم، اقدام به ترجمه چک‌لیست خطر سوگیری، ابزار خطر سوگیری جهت ارزیابی کیفیت مطالعات با تأکید بر پایایی یا خطای اندازه‌گیری، چک‌لیست طراحی مطالعه و راهنمای گزارش، به زبان فارسی کنند. همچنین در وبسایت‌های مجلات پژوهشی، یک فرمت اختصاصی برای مطالعات روان‌سنجی ابزار براساس چک‌لیست‌های کاسمین به‌عنوان راهنمای نویسندگان و داوران مقالات روان‌سنجی، تنظیم و بارگذاری شود.

### ملاحظات اخلاقی

#### پیروی از اصول اخلاق پژوهش

مطالعه مروری حاضر با تأییدیه اخلاق در دانشگاه علوم پزشکی کاشان انجام شده است (کد اخلاق: IR.KAUMS.NUHEPM. REC.1401.047).

21. Principal Axis Factoring
22. Common Factor Analysis
23. Orthogonal Rotation
24. Oblique Rotation



## References

- [1] Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018; 27(5):1147-57. [DOI:10.1007/s11136-018-1798-3] [PMID] [PMCID]
- [2] Griffiths C, Armstrong-James L, White P, Rumsey N, Pleat J, Harcourt D. A systematic review of patient reported outcome measures (PROMs) used in child and adolescent burn research. *Burns.* 2015; 41(2):212-24. [DOI:10.1016/j.burns.2014.07.018] [PMID]
- [3] Ritmala-Castren M, Lakanmaa RL, Virtanen I, Leino-Kilpi H. Evaluating adult patients' sleep: An integrative literature review in critical care. *Scand J Caring Sci.* 2014; 28(3):435-48. [DOI:10.1111/scs.12072] [PMID]
- [4] Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol.* 2006; 6:2. [DOI:10.1186/1471-2288-6-2] [PMID] [PMCID]
- [5] Mokkink LB, Terwee CB, Stratford PW, Alonso J, Patrick DL, Riphagen I, et al. Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual Life Res.* 2009; 18(3):313-33. [DOI:10.1007/s11136-009-9451-9] [PMID]
- [6] Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res.* 2012; 21(4):651-7. [DOI:10.1007/s11136-011-9960-1] [PMID] [PMCID]
- [7] Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol.* 2010; 10:22. [DOI:10.1186/1471-2288-10-22] [PMID] [PMCID]
- [8] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual Life Res.* 2010; 19(4):539-49. [DOI:10.1007/s11136-010-9606-8] [PMID] [PMCID]
- [9] Haywood KL, Garratt AM, Fitzpatrick R. Quality of life in older people: a structured review of generic self-assessed health instruments. *Qual Life Res.* 2005; 14(7):1651-68. [DOI:10.1007/s11136-005-1743-0] [PMID]
- [10] Wind H, Goutteborge V, Kuijjer PP, Frings-Dresen MH. Assessment of functional capacity of the musculoskeletal system in the context of work, daily living, and sport: A systematic review. *J Occup Rehabil.* 2005; 15(2):253-72. [DOI:10.1007/s10926-005-1223-y] [PMID]
- [11] Marinus J, Ramaker C, van Hilten JJ, Stiggelbout AM. 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. [DOI:10.1136/jnnp.72.2.241] [PMID] [PMCID]
- [12] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010; 63(7):737-45. [DOI:10.1016/j.jclinepi.2010.02.006] [PMID]
- [13] Tate RL. Measuring outcomes and monitoring progress in the era of evidence-based clinical practice. *Brain Impair.* 2019; 20(3):276-88. [DOI:10.1017/BrImp.2019.28]
- [14] Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res.* 2009; 18(8):1115-23. [DOI:10.1007/s11136-009-9528-5] [PMID] [PMCID]
- [15] Beattie M, Lauder W, Atherton I, Murphy DJ. Instruments to measure patient experience of health care quality in hospitals: A systematic review protocol. *Syst Rev.* 2014; 3:4. [DOI:10.1186/2046-4053-3-4] [PMID] [PMCID]
- [16] Lenz ER. *Measurement in nursing and health research.* New York: Springer Publishing Company; 2010. [Link]
- [17] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007; 60(1):34-42. [DOI:10.1016/j.jclinepi.2006.03.012] [PMID]
- [18] Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res.* 2018; 27(5):1171-9. [DOI:10.1007/s11136-017-1765-4] [PMID] [PMCID]
- [19] COSMIN. Checklists for Assessing Study Qualities [Internet] 2022 [Updated 2023 April 18]. [Available from: [Link]
- [20] Mokkink LB, Boers M, van der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: A Delphi study. *BMC Med Res Methodol.* 2020; 20(1):293. [DOI:10.1186/s12874-020-01179-5] [PMID] [PMCID]
- [21] Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, de Vet HC, et al. COSMIN study design checklist for patient-reported outcome measurement instruments. *Gut.* 2020; 70:1-32. [DOI:10.1136/gutjnl-2020-320729] [PMID]
- [22] Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions.* 2011 [Updated 2011 March 11]. Available from: [Link]
- [23] Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Qual Life Res.* 2018; 27(5):1159-70. [DOI:10.1007/s11136-018-1829-0] [PMID] [PMCID]
- [24] Qin S, Nelson L, McLeod L, Eremenco S, Coons SJ. Assessing test-retest reliability of patient-reported outcome measures using intraclass correlation coefficients: Recommendations for selecting and documenting the analytical formula. *Qual Life*

- Res. 2019; 28(4):1029-33. [DOI:10.1007/s11136-018-2076-0] [PMID] [PMCID]
- [25] 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; 30(8):2197-218. [DOI:10.1007/s11136-021-02822-4] [PMID]
- [26] Bull C, Byrnes J, Hettiarachchi R, Downes M. A systematic review of the validity and reliability of patient-reported experience measures. *Health Serv Res.* 2019; 54(5):1023-35. [DOI:10.1111/1475-6773.13187] [PMID] [PMCID]
- [27] McKenna SP, Heaney A. Setting and maintaining standards for patient-reported outcome measures: Can we rely on the COSMIN checklists? *J Med Econ.* 2021; 24(1):502-11. [DOI:10.1080/13696998.2021.1907092] [PMID]
- [28] Szekeres M. Clinical relevance commentary in response to: The validity and clinical utility of the Disabilities of the Arm Shoulder and Hand questionnaire for hand injuries in developing country contexts: A systematic review. *J Hand Ther.* 2018; 31(1):91-2. [DOI:10.1016/j.jht.2017.12.002] [PMID]
- [29] Kwok EYL, Rosenbaum P, Thomas-Stonell N, Cunningham BJ. Strengths and challenges of the COSMIN tools in outcome measures appraisal: A case example for speech-language therapy. *Int J Lang Commun Disord.* 2021; 56(2):313-29. [DOI:10.1111/1460-6984.12603] [PMID]
- [30] de Vet HC, Adèr HJ, Terwee CB, Pouwer F. Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Qual Life Res.* 2005; 14(5):1203-18. [DOI:10.1007/s11136-004-5742-3] [PMID]