



# Approaches of measuring the minimum clinically important differences in clinical trial studies in the field of health and treatment: A narrative review of concepts and methods

Nahid Bigdeli<sup>1</sup>, Ali Yalfani<sup>2\*</sup>

1. Ph.D. candidate, Department of Exercise Rehabilitation, Faculty of Sport Sciences, Bu-Ali Sina University, Hamedan, Iran

2. Professor, Department of Exercise Rehabilitation, Faculty of Sport Sciences, Bu-Ali Sina University, Hamedan, Iran

Received 24 January 2024; Accepted 29 February 2024

## Keywords

Minimal clinically important difference  
Clinical trial  
Anchor-based  
Distribution-based

## Abstract

**Background and aim:** In recent years, statistical significance alone cannot determine the impact of interventions in RCT studies, and the minimum clinically important difference (MCID) is used to check the real value of an intervention to enter the treatment and health system. The present study is a narrative review of the concepts and methods of MCID calculation in clinical trial studies to understand this diagnostic approach better.

**Methods:** Researchers searched Google Scholar, Web of Science, and PubMed by combining the keywords "MCID", "distribution-based approach", "anchor-based approach" and "minimum clinically important difference" in English. , and reviewed more than a thousand studies related to the determination and calculation of MCID, and finally the general concept of MCID and related methods were identified and the basic assumptions, statistical biases, and shortcomings of each method were examined in detail.

**Finding:** Two acceptable methods compared to other methods were identified, distribution-based and anchor-based. Also, the general limitations in the exact determination of an identified MCID were: the number and difference of scores obtained from different methods of MCID calculation, the loss of the patient's perspective, and the relationship between the change scores before and after treatment.

**Conclusion:** The purpose of MCID calculation is to determine the most optimal treatment method both in terms of time and cost for patients. In recent years, the importance of this issue seems more and more; Therefore, researchers and specialists in clinical fields can use these methods cautiously to measure the performance of their interventions in the work and study environment.

\* Corresponding Author: Tel: 09183155478

✉ Email: ali\_yalfani@yahoo.com

ORCID: 0000-0003-3696-7208

## Extended Abstract

### Introduction

One of the ways to measure the effect of intervention in clinical trial studies is to use statistical tests (significance). In clinical trial studies, sometimes the treatment is clinically effective. Still, statistically, no significant difference is found (for example, due to the small sample size), or sometimes, this subject is the opposite (1). One of the ways to measure the effect of intervention in clinical trial studies is to use statistical tests (significance). In clinical trial studies, sometimes the treatment is clinically effective. Still, statistically, no significant difference is found (for example, due to the small sample size), or sometimes this subject is the opposite. A statistically significant difference is seen, but there was no significant change in the patient's clinical condition before and after the intervention. The patient still suffers from the problem. Therefore, entering it into the health and medical care system is a good use of time and cost (2). In clinical research, it cannot be said with confidence that statistically significant changes mean treatment or improvement of the desired disorder or disease, so minimum clinically important difference (MCID) was introduced by Jeashka et al. in 1989 to solve the shortcomings of statistically significant differences. They reasoned that although statistically significant changes often occur during tools that measure changes before and after an intervention, clinical changes are of low importance in some cases, and entering them into the health and medical care system is worthless (3-5).

### Finding approaches to the MCID

#### *Distribution based approach*

The distribution-based approach's calculation methods to determine the MCID score are based on statistical characteristics and sample size. They are used when patients are unavailable to be surveyed and describe their conditions regarding the intervention (4,6).

#### *The MDC*

The minimum detectable change index is related to the standard error of measurement (SEM), called Minimum detectable change (MDC). This index alone does not provide information about the clinical significance of a change because it is not dependent on clinical changes and is valued by the MCID. In fact, the MDC indicates to the researcher the minimum

detectable change that can be obtained above the level of measurement error with a given confidence level (usually a 95% confidence level)(5). The studies in this field acknowledge that an MCID has value and validity when its score is higher than MDC or at least equal to it (6,7). The following formula shows how to calculate this index:

$$MDC_{95\%} = 1.96 \times \sqrt{2} \times SEM$$

#### *Calculate the MCID based on the SEM*

SEM is the standard deviation of the measurement error in an assessment or test. This index, which is related to error variance and the Intraclass Correlation Coefficient (ICC), is a measurement tool that shows how many of the changes that occurred in a group after the intervention can be caused by measurement error (8,9). To calculate this index, we need the standard deviation or SD of the pre-test or baseline to determine how much error we may encounter based on the reliability coefficient of the test. To calculate the MDC in the previous paragraph, one must first calculate the SEM using the following formula:

$$SEM = SD_{pre\ or\ baseline} \times \sqrt{1 - ICC}$$

In the preceding formula, ICC is related to the reliability coefficient of the devices, tools, or tests used to measure the desired variable. We also use the following formula to calculate MCID based on SEM:

$$MCID = X \times SD_{pre\ or\ baseline} [\sqrt{(1 - ICC)}]$$

In the prior formula, we chose X according to the study's confidence level (1=small, 1/96=medium, 2/77=large). However, in clinical studies, X is usually 1.96 (95% confidence level).

#### *MCID Calculation based on the RCI*

Another concept we are familiar with about MCID is the reliable change index (RCI). If the RCI is greater than 1.96, the change in the patient is considered an actual change with 95% confidence (9). This method suits studies with a high confidence level and a high-reliability coefficient in tools and tests(9-12). Therefore, finally, to calculate MCID based on RCI, we use the following formula, where X is usually equal to 1.96:

$$MCID = X \times SD_{pre\ or\ baseline} \{ \sqrt{[2 \times (1 - ICC)]} \}$$

#### *Anchor-based approach*

The anchor-based approach relates the change in the outcome score to an external criterion that shows the patient's view of improvement. The external criterion can be a questionnaire or a rating scale that assesses the patient's well-being or symptom level. The anchor-based approach is helpful for the clinician who wants to track the patient's improvement and reaction to the intervention (4-6,8)

#### ***The approach based on sensitivity and specificity***

In this approach, we choose a score as MCID with the best-discriminating power among patient groups (the score that creates the highest sensitivity and specificity). About MCID, sensitivity is the proportion of patients whose improvement was correctly diagnosed (actual positive changes). Specificity is also the proportion of patients who have no change, and this lack of change is correctly recognized (actual negative changes). A sensitivity score of 1 indicates that all actual positive cases are detected, and a specificity score of 1 indicates that all actual negative cases are detected. In this method, like other methods, there is an error caused by wrong reporting of patient improvement, which is false negative and false positive (2).

The ROC<sup>1</sup> curve is used to identify the PRO score with equal sensitivity and specificity to distinguish between patients with "improvement" and "no improvement" in the method based on sensitivity and specificity. A ROC curve is a graph of sensitivity (TPR<sup>2</sup>) versus specificity-1 (FPR<sup>3</sup>) as the cutoff value moves from 0 to 1. In the ROC curve, a model with high sensitivity and high specificity will approach the upper left corner of the plot, and a model with low sensitivity and specificity will have a curve close to the bisector line (11). ROC curve models are well displayed in Figure 3. From right to left, the curve shows the best effect (in intervention) or detection (in test measurement), and the space under the curve is 100%. As we get closer to the left side of the arrow, the curve is closer to the bisector line and eventually crosses it, and the space under the curve gets smaller (24).

In this case, we have a false positive or zero in the center of the graph, a real positive graph (green line) in the upper corner, and a real negative graph (red line) in the lower corner. The probability range of the area under the curve (AUC) varies from 0.5 to 1. 0.5 indicates the ability to diagnose due to chance, and 1 indicates

the ability to diagnose improvement in all patients correctly, and of course, the area of 0.8 to 0.9 is good. An area of 0.7 to 0.8 is considered relatively good. AUC indicates change or lack of change in patients, a false negative in the center, and a true negative in the corner of the graph (12).

#### ***Conclusion***

In the present study, the authors explained the concepts of minimum clinically important difference or MCID and the most important approaches to obtain it. The most useful implication of MCID is that the treatment is effective and the physician and patient are aware of its impact on the patient's life. Future researchers should better understand the factors influencing MCID scores that are specific to the study population. Use of the MCID score in both the clinical and research settings is prudent until consensus can be reached on the calculation to address methodological, population, and pretest dependence limitations.

<sup>1</sup> Receiver Operating Characteristic

<sup>2</sup> True positive rate

<sup>3</sup> False positive rat



# رویکردهای اندازه‌گیری حداقل تفاوت مهم بالینی در مطالعات کارآزمایی بالینی حوزه سلامت و درمان: یک مرور روایتی بر مفاهیم و روش‌ها

ناهید بیگدلی<sup>۱</sup>، علی یلفانی<sup>۲\*</sup>

۱- دانشجوی دکتری، گروه توانبخشی ورزشی، دانشکده علوم ورزشی، دانشگاه بوعلی سینا، همدان  
۲- استاد، گروه توانبخشی ورزشی و حرکات اصلاحی، دانشکده علوم ورزشی، دانشگاه بوعلی سینا، همدان

مقاله مروری

دریافت ۰۶ بهمن ۱۴۰۲؛ پذیرش ۱۰ اسفند ۱۴۰۲

## واژگان کلیدی

حداقل تفاوت مهم بالینی

کارآزمایی بالینی

مبتنی بر لنگر،

مبتنی بر توزیع

## چکیده

**زمینه و هدف:** در برخی مطالعات سال‌های اخیر به نظر می‌رسد که گاهی معنی‌داری آماری در مطالعات RCT کافی نیست و برای بررسی ارزش واقعی یک مداخله جهت ورود به سیستم درمان و بهداشت، از حداقل تفاوت مهم بالینی (MCID) استفاده می‌گردد. مطالعه حاضر یک مرور روایتی بر مفاهیم و روش‌های محاسبه MCID در مطالعات کارآزمایی بالینی جهت شناخت بهتر این رویکرد تشخیصی است.

**روش بررسی:** محققین با جستجو در گوگل اسکالر، Web of science و Pub med با ترکیب واژه‌های کلیدی «MCID»، «رویکرد مبتنی بر توزیع»، «رویکرد مبتنی بر لنگر»، «حداقل تفاوت مهم بالینی» به زبان انگلیسی، بیش از هزار مطالعه پیدا کردند که از این بین مقالات مرتبط با تعیین و محاسبه MCID مطالعه و در نهایت مفهوم کلی MCID و روش‌های مرتبط با آن شناسایی شد و مفروضات اساسی، سوگیری‌های آماری و کاستی‌های هر روش به تفصیل مورد بررسی قرار گرفت.

**یافته‌ها:** دو روش مبتنی بر توزیع و مبتنی بر لنگر نسبت به سایر روش‌ها قابل قبول شناخته شدند. همچنین محدودیت‌های کلی در تعیین دقیق یک MCID شناسایی شده عبارت بود از: تعدد و تفاوت نمره‌های به‌دست‌آمده از روش‌های مختلف محاسبه MCID، ازدست‌دادن دیدگاه بیمار و رابطه بین نمرات تغییر قبل و بعد از درمان.

**نتیجه‌گیری:** هدف از محاسبه MCID تعیین بهینه‌ترین روش درمان چه از نظر وقت و چه از نظر هزینه برای بیماران است. در سال‌های اخیر اهمیت این موضوع بیش‌ازپیش به نظر می‌رسد؛ بنابراین محققان و متخصصان رشته‌های بالینی می‌توانند با احتیاط نسبت به محدودیت‌های موجود در ارتباط با این روش‌ها، برای سنجش عملکرد مداخلات خود در فضای کار و مطالعه از آنها استفاده کنند.

## مقدمه

همکارانش در سال ۱۹۸۹ به عنوان راهی برای رفع نواقص تفاوت معنی‌دار آماری معرفی شد، یعنی کمترین تغییر مهم در یک متغیر یا دامنه مثل تعادل، کیفیت زندگی، درد و ... . استدلال آنها این بود که اگرچه از نظر آماری تغییرات قابل توجهی اغلب در طول استفاده از ابزارهایی که تغییرات را پس از مداخله اندازه‌گیری می‌کنند رخ می‌دهد، اما در برخی موارد تغییرات معنی‌دار از اهمیت بالینی پایینی برخوردارند و به لحاظ وقت و هزینه ارزش ورود به سیستم درمانی را ندارند (۳، ۴). کوپای و همکاران (۲۰۰۷) در ارتباط با تفاوت‌های آماری در مطالعات، اظهار داشتند: «یک تفاوت می‌تواند از نظر آماری و بر اساس یک مقدار عددی ساده، معنی‌دار باشد، اما در عین حال ممکن است برای سلامت یا کیفیت زندگی بیماران مبتلا به یک بیماری خاص، اهمیت کمی داشته باشد یا هیچ اهمیتی نداشته باشد. علاوه بر این، حجم نمونه مورد آزمایش اغلب به اهمیت آماری یک متغیر خاص کمک می‌کند، به طوری که ممکن است جزئیات به ظاهر بی‌اهمیت آشکار و معنی‌دار شود» (۵). یک تفاوت مهم کلینیکالی زمانی پدیدار می‌شود که نتایج مداخله از نظر بیمار ارزشمند و معنادار تلقی شود و در صورت لزوم بیمار علاقه‌مند به تکرار مجدد آن باشد (۳). لازم به ذکر است که MCID تنها راه ممکن برای ارزیابی نتایج کارآزمایی‌های بالینی نیست و سایر رویکردها مانند حجم نمونه و اندازه اثر نیز ممکن است اطلاعات ارزشمندی را در تفسیر کارآزمایی‌های بالینی ارائه دهند، اما اطلاعاتی که MCID در رابطه با تغییرات بالینی می‌دهد موجب می‌شود که محقق نتایج مطمئن‌تر و قوی‌تری استنباط کند؛ از گذشته تا به امروز مطالعات بالینی بسیاری در سراسر دنیا در حال استفاده از این روش هستند؛ لیکن بنظر می‌رسد که این رویکرد به دلیل عدم شناخت در مطالعات علمی ایران کمتر مورد استفاده قرار گرفته است. لذا هدف از مطالعه حاضر، مروری روایتی بر مفاهیم و روش‌های اندازه‌گیری MCID در تحقیقات کارآزمایی بالینی می‌باشد.

## روش‌ها

مطالعه حاضر یک مرور روایتی است که در آن به بررسی مفهوم «حداقل تفاوت مهم بالینی»، علت و چگونگی استفاده از آن در تحقیقات حوزه بالینی پرداخته می‌شود. محققین مطالعه حاضر با جستجو در پایگاه‌های گوگل

مطالعات کارآزمایی بالینی تحقیقاتی هستند که در رشته‌های پزشکی، پیراپزشکی، علوم ورزشی، روانشناسی و حوزه‌های مرتبط با سلامت کاربرد دارند و هدف آن‌ها، تعیین اثربخشی و ایمنی یک آزمایش یا پروتکل درمانی جدید است که در آن از یک یا چند گروه مداخله و یا کنترل استفاده می‌شود. نتایج مستخرج از مطالعات کارآزمایی بالینی می‌توانند به میزان قابل توجهی شیوه‌های مراقبت بالینی از بیماران را تغییر دهند و در ارتقای سلامت آنها تأثیرگذار باشند (۱). افزایش هزینه مراقبت‌های بهداشت و درمان، مسئولیت پاسخگویی برای ارائه مداخلات درمانی آگاهانه و موثر را الزامی می‌کند. لذا میزان تأثیر یک مداخله و یا درمان در یک مطالعه کارآزمایی بالینی از موارد بسیار مهمی است که توجه مسئولین سلامت و بهداشت را به خود معطوف می‌دارد. سیستم‌های بهداشت و درمان در سرتاسر جهان از مداخلاتی استقبال می‌کنند که با کمترین هزینه برای بیماران، بیشترین کارایی و اثربخشی را به ارمغان بیاورد؛ به عبارتی دیگر، بیماران بیشترین سود را ببرند (۲). یکی از راه‌های سنجش اثر مداخله در مطالعات کارآزمایی بالینی، استفاده از آزمون‌های آماری (معنی‌داری) است. در دنیای آمار، تفاوت معنادار صرفاً تفاوتی است که بعید است، ناشی از تصادف یا اتفاق باشد و بر اساس مبنای ریاضی است. در مطالعات کارآزمایی بالینی گاهی درمان به لحاظ کلینیکالی نتیجه بخش است اما از نظر آماری، تفاوت معنی‌داری یافت نمی‌شود (مثلاً به دلیل مقدار کم حجم نمونه)، یا حتی بالعکس به لحاظ آماری تفاوت معنی‌داری یافت شده است اما وضعیت بیمار به لحاظ کلینیکالی، قبل و بعد از درمان تغییر چشم‌گیری نکرده است و بیمار همچنان از مشکل موردنظر رنج می‌برد (مثلاً تعداد بالای حجم نمونه و حساسیت آزمون آماری به تغییرات، موجب معنی‌داری شده است). بنابراین ورود این نوع درمان به سیستم بهداشتی، اتلاف وقت و هزینه است (۳). لذا در تحقیقات بالینی نمی‌توان با اطمینان خاطر گفت که تغییرات معنی‌دار آماری به معنای درمان و یا بهبود عارضه و یا بیماری موردنظر است. بنابراین «حداقل تفاوت مهم کلینیکالی»<sup>۵</sup> (MCID) برای اولین بار توسط پیشکا و

<sup>5</sup> Minimum clinically important difference

دهد که آیا تغییرات آماری تصادفی بوده است یا صرفاً وجود مداخله درمانی موجب این تغییرات بوده است (۸). در مقابل، این رویکرد نقاط ضعفی نیز دارد. یکی از نقاط ضعف روش‌های مبتنی بر توزیع این است که به نظرسنجی از بیماران در رابطه با تغییرات به وجود آمده نمی‌پردازد و دیگری نیز مرتبط با حجم نمونه است، مخصوصاً اگر حجم نمونه بالا باشد (۳، ۸)؛ مثلاً در یک مطالعه با حجم نمونه بالا، حتی اگر تغییرات واقعی نباشند، باز هم می‌توان نمره MCID را با تجزیه و تحلیل آماری استخراج کرد، و البته که در این مواقع نمی‌توان به این نمره اعتماد کاملی داشت؛ لذا برای حصول اطمینان بیشتر، احتمال این اتفاق را با شاخصی به نام MDC یا "حداقل تغییر قابل تشخیص" پایین می‌آوریم (۹).

### شاخص MDC

شاخص MDC نشان‌دهنده «خطای استاندارد اندازه‌گیری»<sup>۹</sup> یا SEM<sup>۱۰</sup> است که به آن «حداقل تغییر قابل تشخیص» یا «کوچکترین تغییر قابل تشخیص» هم گفته می‌شود. این شاخص به تنهایی نمیتواند اهمیت بالینی یک تغییر را نشان دهد، زیرا به تغییرات بالینی وابسته نیست و فقط در کنار شاخص MCID معنا پیدا میکند. در واقع، MDC حداقل تغییری که بیشتر از سطح خطای اندازه‌گیری با سطح اطمینان ۹۵ درصد به دست بیاید را به محقق نشان می‌دهد. معمولاً این شاخص در مطالعات به شکل ۹۵٪ MDC نشان داده می‌شود (۱۰)؛ بنابراین ارزش نمره MCID می‌تواند کمتر و یا بیشتر از MDC باشد. مطالعاتی که در این زمینه انجام شده است اذعان دارند که یک MCID زمانی دارای ارزش و اعتبار است که نمره آن بالاتر از MDC و یا حداقل به اندازه آن باشد (۵، ۹، ۱۱)؛ به بیانی دیگر MDC مانند یک نمره مرجع عمل می‌کند که محقق زمانی می‌تواند با اطمینان بگوید که تغییرات مشاهده شده بر اثر مداخلات درمانی است و نه تغییرات تصادفی، که نمره MCID متغیرهای آن مطالعه، بالاتر از نمره MDC باشد. برای محاسبه‌ی این شاخص اول SEM را محاسبه می‌کنیم که نیاز به انحراف استاندارد یا SD<sup>۱۱</sup> پیش آزمون داریم تا بتوان تشخیص داد بر اساس ضریب

اسکالر، Web of science و Pub med با ترکیب واژه‌های کلیدی «MCID»، «رویکرد مبتنی بر توزیع»، «رویکرد مبتنی بر لنگر» و «حداقل تفاوت مهم بالینی» به زبان انگلیسی بیش از هزار مطالعه مرتبط با تعیین و محاسبه MCID را پیدا و بررسی کردند و در نهایت موضوع مطالعه پس از مطالعه‌ی عناوین و چکیده‌های مختلف در ارتباط با MCID در مطالعاتی که بین سال‌های ۱۹۸۹ تا ۲۰۲۳ انجام شدند انتخاب شد. همچنین مقالات غیرمرتبط با موضوع مطالعه و رویکردهای مبتنی بر لنگر<sup>۶</sup> و مبتنی بر توزیع<sup>۷</sup> از روند مطالعه خارج شدند. معیارهای ورود به مطالعه حاضر، شامل دامنه‌ی مطالعات مرور مفهومی و مطالعات کارآزمایی بالینی با روش اندازه‌گیری MCID بود.

### روش‌های تعیین MCID

با بررسی این مطالعات مشخص شد که رویکردهای مورد استفاده برای تعیین یک MCID ثابت نیستند. یک محقق به نام ولز راجع به روش‌های محاسبه و اندازه‌گیری MCID تحقیقات گسترده‌ای انجام داد و نه رویکرد مختلف را معرفی کرد (۶). پس از مدتی محقق دیگر به نام لیزر این رویکردها را به سه گروه تقسیم‌بندی کرد: رویکرد مبتنی بر توزیع، رویکرد مبتنی بر لنگر یا پیش‌بینی‌کننده، و رویکرد مبتنی بر نظرسنجی<sup>۸</sup> (که در آن نظرات بیماران، پزشکان و متخصصین به شکل مصاحبه جمع‌آوری می‌شود) (۷). محققین در این مطالعه رویکرد مبتنی بر توزیع و مبتنی بر لنگر را مورد بحث قرار خواهند داد.

### یافته‌ها

#### رویکرد مبتنی بر توزیع (Distribution-based)

روش‌های محاسبه رویکرد مبتنی بر توزیع برای تعیین نمره MCID براساس ویژگی‌های آماری و مختص نمونه هستند و همچنین در زمان‌هایی مورد استفاده قرار می‌گیرد که بیماران جهت نظرسنجی و توصیف شرایط خود در رابطه با مداخله، در دسترس نباشند و به هر دلیلی از دسترس خارج شده باشند. رویکرد مبتنی بر توزیع توانایی بالایی در محاسبه تغییرات به وجود آمده فراتر از سطوح تغییرات تصادفی دارد؛ به عبارت دیگر، این رویکرد می‌تواند نشان

<sup>9</sup> Minimal Detectable Change

<sup>10</sup> Standard Error of Measurement

<sup>11</sup> Standard Deviation

<sup>6</sup> Anchor-based or predictive

<sup>7</sup> Distribution-based

<sup>8</sup> Opinion-based



در شکل ۱. ضریب پایایی (ICC) شاخص نائینگهام ۰/۹۷ و SD پیش‌آزمون ۱۲/۱ است که در فرمول جایگذاری می‌شوند. و در نهایت، نمره MDC فعالیت‌های گسترده روزمره زندگی نائینگهام ۴/۹ بدست آمده است و نمره‌ی MCID آن ۶/۱. بنابراین، بالاتر بودن نمره MCID نسبت به MDC نشان دهنده اعتبار آن در تشخیص اثر مداخله است (۱۲). حال اگر نمره MCID کوچکتر از نمره MDC باشد، نتیجه عکس این موضوع خواهد بود. لازم به ذکر است که MDC مرجعی برای هر دو روش مبتنی بر توزیع و لنگر است.

پایایی آزمون، با چه مقدار خطای ناشی از اندازه‌گیری مواجه هستیم. سپس از فرمول زیر استفاده می‌کنیم:

$$SEM = SD_{pre \text{ or } baseline} \times \sqrt{1 - ICC}$$

در فرمول بالا، ICC مرتبط با ضریب پایایی دستگاه، وسیله و یا تستی است که برای سنجش متغیر مورد نظر مورداستفاده قرار گرفته است. فرمول ذیل، نحوه محاسبه این شاخص را نشان می‌دهد:

$$MDC_{95\%} = 1.96 \times \sqrt{2} \times SEM$$

Scale	SD*	ICC	Standard Error of Measurement	MDC <sub>90</sub>	MCID Estimates				
					% Scale Width of the MDC <sub>90</sub>	0.2 SD	% Scale Width of 0.2SD	Anchor-Based MCID	% Scale Width of the Anchor-Based MCID
NEADL total score	12.1	.97	2.1	4.9	8.6	2.4	4.2	6.1	10.6

شکل ۱- بررسی یک نمونه از ارتباط MCID و MDC و ارتباط آن با معنی داری کلینیکالی یک مداخله (وو و همکاران، ۲۰۱۱).

محاسبه می‌کند که با توجه به روایی بالای این وسیله، استاندارد محسوب می‌شود. معمولاً سطح اطمینان ۱ تا ۱/۹۶، در محاسبه MCID برآورد خوبی را به ما می‌دهد (۱۴). برای محاسبه MCID بر اساس SEM از فرمول زیر استفاده می‌کنیم:

$$MCID = X \times SD_{pre \text{ or } baseline} [\sqrt{(1 - ICC)}]$$

در این فرمول X را با توجه به سطح اطمینان مطالعه، انتخاب می‌کنیم (۱ = کوچک، ۱/۹۶ = متوسط، ۲/۷۷ = بزرگ). اگرچه در مطالعات بالینی، X معمولاً ۱/۹۶ (سطح ۹۵ درصد اطمینان) است.

### محاسبه بر اساس انحراف استاندارد

انحراف استاندارد (SD) میزان پراکندگی بین یک گروه از نمرات را در دامنه‌ی بالاتر و پایین‌تر از میانگین نشان می‌دهد. برای برآورد MCID فقط از SD گروه تمرین در پیش آزمون و قبل از هرگونه مداخله درمانی استفاده می‌کنیم. از نظر برخی محققین استفاده از SD ۰/۵ برای برآورد MCID در زمانی که روایی یک دستگاه و یا وسیله ارزیابی بالا نیست بلامانع است. این روش یک برآورد متوسط و تقریبی در رابطه با اثر مداخلات است. به طور

### روش‌های محاسبه MCID بر اساس رویکرد مبتنی بر توزیع

در مطالعه حاضر، از چهار روش: خطای استاندارد اندازه‌گیری (SEM)، انحراف استاندارد (SD)، شاخص تغییرات قابل‌اطمینان<sup>۱۲</sup> (RCI) و اندازه اثر<sup>۱۳</sup> (ES) برای محاسبه MCID استفاده خواهیم کرد. هرکدام از این روش‌ها با توجه به شرایط یک مطالعه انتخاب می‌شوند:

### محاسبه MCID بر اساس خطای استاندارد اندازه‌گیری

SEM یا خطای استاندارد اندازه‌گیری، انحراف استاندارد خطای اندازه‌گیری در یک ارزیابی و یا آزمون می‌باشد. این شاخص مرتبط با واریانس خطا و ضریب پایایی<sup>۱۴</sup> (ICC) یک وسیله اندازه‌گیری است، که نشان می‌دهد چه مقدار از تغییراتی که در یک گروه بعد از مداخله به وجود آمده، می‌تواند ناشی از خطای اندازه‌گیری باشد (۱۳)؛ به طور مثال زمانی که پایایی یک وسیله ارزیابی ۰/۹۷ باشد، احتمال ۰/۰۳ خطا در این وسیله وجود دارد و SEM به طور حدودی، تغییرات ناشی از خطای وسیله و یا تست را

<sup>12</sup>Reliable change index

<sup>13</sup> Effect size

<sup>14</sup> Intraclass Correlation Coefficient

متوسط و ۰/۸ را بزرگ در نظر گرفته است. اندازه اثری که همراه با یک معیار خارجی استفاده می‌شود، پاسخگویی معیار خارجی را مشخص می‌کند. به عنوان مثال، اندازه اثر در پرسشنامه ناتوانی و درد، در بیمارانی که هیچ تغییری را گزارش نمی‌کنند کوچک و در بیمارانی که بهبود زیادی را گزارش می‌کنند، بزرگ می‌باشد (۱۷). برخی مطالعات بیان داشته‌اند که تغییر در امتیازات مربوط به اندازه اثر کوچک را می‌توان به عنوان MCID در نظر گرفت؛ به این معنی که بهتر است برای محاسبه MCID با این روش، SD پیش-آزمون را در ۰/۲ ضرب کنیم (اندازه اثر کوچک)؛ اگرچه برخی از مطالعات برای جلوگیری از ورود نمره‌ی MCID به مرز خطای اندازه‌گیری (۵ درصد خطای آلفا)، اندازه اثر متوسط یعنی ۰/۵ را به جای X قرار می‌دهند. در فرمول زیر، X همان مقدار اندازه اثر دی کوهن می‌باشد.

$$MCID = X \times SD_{Pre \text{ or } baseline}$$

#### تفسیر MCID بر اساس روش مبتنی بر توزیع

همانطور که گفته شد، MCID حداقل تفاوتی است که یک روش درمانی باید از پیش‌آزمون تا پس‌آزمون داشته باشد در غیر این صورت آن روش درمانی نامعتبر شناخته می‌شود. برای تفسیر نتایج MCID باید مقدار آن را با توجه توضیحاتی که داده شد و همچنین مشخصات مطالعه خود و فرمول مناسب آن طبق جدول ۱، برای هر ابزار و روش اندازه‌گیری محاسبه و سپس عدد بدست آمده را با شاخص MDC مقایسه کرد (۵، ۱۱). به طور مثال در مطالعه یلفانی و همکاران ۲۰۲۲، محققین برای بررسی اثر کلینیکالی متغیر IRD انحراف استاندارد پیش‌آزمون گروه مداخله را در این متغیر بدست آوردند؛ سپس براساس آن به ترتیب نمره-ی SEM، MDC و MCID را محاسبه و طبق مقایسه MCID با تفاوت میانگین پیش‌آزمون و پس‌آزمون<sup>۱۵</sup> (نمره  $\Delta$ ) مداخلات اثر مثبتی از خود نشان داده است (۱۸):

مثال اگر پایایی یک ابزار اندازه‌گیری ۰/۷۵ باشد، حداقل تفاوت مهم بالینی ۱/۲ انحراف استاندارد (SD ۰/۵) محسوب می‌شود و به شکل ذیل محاسبه می‌گردد (۱۵)؛ لذا این روش محاسبه MCID را معمولاً مطالعاتی می‌توانند انجام دهند که اعتبار ابزار آن‌ها متوسط است (۱۱، ۱۵).

$$MCID = 0.5 \times SD$$

#### محاسبه بر اساس شاخص تغییرات پایا

مفهوم دیگری که در رابطه با MCID با آن آشنا می‌شویم، شاخص تغییرات پایا یا RCI است، این شاخص به تغییراتی اشاره دارد که بیشتر از تغییرات ناشی از خطای اندازه‌گیری است و نمره آن به صورت زیر به دست می‌آید:

$$RCI = SD_{pre \text{ or } baseline} \times [2 \times (1 - ICC)]$$

اگر RCI بزرگتر از ۱/۹۶ باشد، تغییر در بیمار به عنوان یک تغییر واقعی با اطمینان ۹۵٪ در نظر گرفته می‌شود (۱۶). MCID به خودی خود یک MCID بالقوه را تشکیل نمی‌دهد، اما در ارتباط با روش‌های دیگر در مطالعات MCID استفاده شده است. این روش برای مطالعاتی که سطح اطمینان مطلوب و ضریب پایایی بالایی در ابزار و تست‌ها دارند، مناسب است (۱۶). لذا در نهایت برای محاسبه MCID بر اساس MCID از فرمول زیر استفاده می‌کنیم که معمولاً X برابر با ۱/۹۶ می‌باشد:

$$MCID = X \times SD_{pre \text{ or } baseline} \{ \sqrt{2 \times (1 - ICC)} \}$$

#### محاسبه بر اساس اندازه اثر (ES)

اندازه اثر یا ES، یک معیار استاندارد شده از تغییر است که با تقسیم تفاوت امتیازات پس‌آزمون و پیش‌آزمون بر انحراف معیار پیش‌آزمون به دست می‌آید.

$$ES_{Cohen's D} = \frac{Score_{mean \text{ posttest}} - Score_{mean \text{ pretest}}}{SD_{pretest}}$$

مقدار اندازه اثر نشان می‌دهد که به لحاظ آماری چقدر تغییر از پیش‌آزمون تا پس‌آزمون به وجود آمده است. معروف‌ترین شاخص نشان‌دهنده اندازه اثر، شاخص دی کوهن است. این شاخص، اندازه اثر ۰/۲ را کوچک، ۰/۵ را

<sup>15</sup> Mean difference



Variable	SD <sub>pre</sub> intervention	Δ Score	SEM	MDC <sub>95%</sub>	MCID
IRD (mm)	9.64	28.58	5.10	14.09	14.35

$$\Delta \text{ Score} \geq \text{MCID} \geq \text{MDC}$$

شکل ۲- بررسی MCID و مقایسه آن با نمره دلتا در مطالعه RCT یلفانی و همکاران (۲۰۲۲)

باشد تا بتوان با اطمینان گفت که مداخله بر متغیر مورد نظر مؤثر بوده است (۱۱، ۱۸). لذا این شرایط در شکل ۲ وجود دارد که نشان می‌دهد مداخله اثربخشی مثبتی داشته است.

بر اساس مطالعات قبلی، دو شرط مهم در بررسی MCID یک مطالعه وجود دارد. اولاً، یک MCID زمانی معتبر است که مقدار آن حداقل برابر یا بیشتر از مقدار MDC باشد. ثانیاً، اختلاف میانگین (نمره Δ) گروه مداخله در پیش‌آزمون و پس‌آزمون باید برابر و یا بیشتر از MCID

جدول ۱- روش‌های محاسبه MCID براساس رویکرد مبتنی بر توزیع

$\text{SEM} = \text{SD}_{\text{pre or baseline}} \times \sqrt{1 - \text{ICC}}$ $\text{MDC}_{95\%} = 1.96 \times \sqrt{2} \times \text{SEM}$	
MCID <sub>(SEM)</sub>	$\text{MCID} = X \times \text{SD}_{\text{pre or baseline}} [\sqrt{(1 - \text{ICC})}]$ <p>در خطای استاندارد اندازه‌گیری، X برای اثر کوچک ۱، برای اثر متوسط ۱/۹۶ و برای اثر بزرگ ۲/۷۷ است که معمولاً از ۱/۹۶ استفاده می‌شود.</p>
MCID <sub>(0.5 × SD)</sub>	$\text{MCID} = 0.5 \times \text{SD} = 1\text{SEM}$ <p>۰/۵ × SD برابر با اندازه اثر ۰/۵ است که به طور کلی در شاخص دی کوهن یک اثر متوسط در نظر گرفته می‌شود.</p>
MCID <sub>(RCI)</sub>	$\text{RCI} = \text{SD}_{\text{pre or baseline}} \times [2 \times (1 - \text{ICC})]$ $\text{MCID} = X \times \text{SD}_{\text{pre or baseline}} \{ \sqrt{[2 \times (1 - \text{ICC})]} \}$ <p>در شاخص تغییرات قابل اطمینان، X معمولاً ۱/۹۶ هست (۹۵٪ اطمینان).</p>
MCID <sub>(ES)</sub>	$\text{MCID} = X \times \text{SD}_{\text{Pre or baseline}}$ <p>با توجه به شاخص دی کوهن، اندازه اثر ۰/۲، ۰/۵ و ۰/۸ به ترتیب تغییرات کوچک، متوسط و بزرگ را تعریف کرده‌اند.</p>
$\text{Mean difference (Md) or } \Delta \text{ Score} = (\text{Mean post} - \text{Mean pre}) \text{ in intervention group}$ $\bar{\Delta} = \sum \Delta 1 + \Delta 2 + \Delta 3 + \dots = \text{MCID}$	

SEM: Standard error measurement; MDC: Minimal detectable change; SD: Standard deviation; MCID: Minimal clinically important difference; RCI: Reliable change index; ES: Effect size.

گیل، رولاند موریس، مقیاس‌های درد، فرم‌های نتایج پزشکی، پرسشنامه‌های روانشناسی، پرسشنامه کیفیت زندگی و ... از جمله ابزار و مواردی هستند که صحت و سقم آن‌ها توسط یک معیار خارجی که به آن لنگر می‌گویند (مانند شاخص‌های GRC یا CGI-C) مشخص می‌شود (۷-۱۱).

این رویکرد می‌تواند روشی برای تعیین کوچکترین تغییر در اندازه‌گیری "نتایج گزارش شده توسط بیمار" یا

### رویکرد مبتنی بر لنگر (Anchor-based)

یکی دیگر از رویکردهایی که برای محاسبه MCID به کار می‌رود، رویکرد مبتنی بر لنگر است. در این روش تغییرات نمرات قبل و بعد پیامدهای اولیه و ثانویه بیماری (متغیرهای مطالعه)، به یک معیار یا لنگر خارجی معتبر مرتبط می‌شود، که دیدگاه بیمار را در رابطه با بهبودی خود مشخص کرده و ارزیابی می‌کند که کدام تغییرات در اندازه‌گیری پیامدهای تعریف شده، با لنگر مطابقت دارد. پرسشنامه‌های مختلف مثل شاخص ناتوانی آسوستری، مک

باشید که برای مقایسه PRO ها با معیار خارجی یا لنگر فقط یک سوال لنگر تعریف می‌شود تا به صورت کلی نظر بیمار در رابطه با وضعیتش پس از دریافت مداخله سنجیده و سپس با میانگین تغییرات PRO مقایسه شود. مثال زیر یک سؤال لنگر مرتبط با کمردرد را نشان می‌دهد:

«لطفاً وضعیت کمردرد خود را از زمان شروع درمان تا به اکنون مشخص کنید (پس از مداخله، کمردرد شما چقدر تغییر کرده است؟)».

باید توجه کرد که سؤال لنگر می‌تواند به‌گونه‌ای طراحی شود که تعداد گزینه‌های مختلفی جهت پاسخ در اختیار بیمار قرار دهد و پاسخ‌ها به‌صورت رتبه‌ای طراحی می‌شوند؛ مثل «بدتر»، «بدون تغییر» یا «بهتر». معمولاً تعداد پاسخ‌های رتبه‌ای همراه با سؤال لنگر، متناسب با پیامد یا ابزار ارزیابی است که برای آن اعمال می‌شود. به طور قراردادی، تعداد گزینه‌های پاسخ به سؤالات لنگر، فرد هستند و تمرکز آنها بر عدم تغییر در نتیجه است (پاسخ‌های عدم نتیجه بیشتر از پاسخ‌های دریافت نتیجه است) که اغلب پاسخ‌ها از ۳ تا ۷ گزینه متغیر می‌باشد (جدول ۲ و ۳).

جدول ۲- شاخص GRC، شناخته شده به عنوان لنگر در میزان تغییرات مداخلات درمانی؛ پیشنهاد شده توسط پیشکا و همکاران (۱۹۸۹).

سؤال: لطفاً وضعیت خود را از زمان شروع درمان تا به اکنون توصیف کنید.	امتیاز شاخص
تغییرات	
به طور کامل بهتر	+۷
خیلی بهتر	+۶
کمی بیشتر از متوسط بهتر	+۵
به طور متوسط بهتر	+۴
تا حدودی بهتر	+۳
کمی بهتر	+۲
خیلی خیلی جزئی بهتر (در حد بدون تغییر)	+۱
بدون تغییر	۰

PRO<sup>۱۶</sup> باشد. این گزارش‌ها روشی ساختار یافته و قابل تکرار برای به دست آوردن اطلاعات بیمار ارائه می‌کنند و البته تحت تأثیر نظر و تفسیر شخص ثالث قرار نمی‌گیرند. این روش اغلب در مطالعاتی که ممکن است سوگیری بین ارزیابان و درمانگران به وجود بیاید، قابل اعتماد است (۱۰). PROها به طور فزاینده‌ای برای ارزیابی نتایج مانند کیفیت زندگی مرتبط با سلامت عمومی، کیفیت زندگی خاص یک بیماری، یا شناخت علائم بیماری مورد استفاده قرار می‌گیرند. رویکرد مبتنی بر لنگر از نظر بالینی برای پزشک یا توانبخشی که می‌خواهد در همان دوره‌ی مراقبت از بیمار متوجه تغییرات مرتبط با مداخله شود، مناسب است (۴). سه دلیل عمده برای استفاده از PROها در تحقیقات بالینی وجود دارد: اول اینکه، بعضی از علائم باید به طور مستقیم از بیمار پرسیده شود و در غیراینصورت راه دیگری برای متوجه شدن آن مشکل وجود ندارد، مثل دردهای گوارشی و یا شکمی؛ دوم، اندازه‌گیری‌های بالینی همیشه با نتایج گزارش شده توسط بیمار مطابقت ندارند؛ گاهی ممکن است بهبودی در یک متغیر موجب بهبودی در متغیر دیگر (مثل درد و یا ناتوانی) نشود (۱۹). به عنوان مثال، ممکن است ارزیابی یک بیمار پس از تمرینات درمانی در ناحیه کمری، بازگشت مناسبی در حس عمقی مفاصل و عضلات آن ناحیه گزارش کند، اما بیمار همچنان از کمردرد شکایت داشته باشد. سوم اینکه، بیمار ممکن است تنها منبع اطلاعات ما باشد، یعنی برخی از اثرات درمانی را فقط بیمار می‌تواند تجربه کند و به طور دقیق قابل ارزیابی نیستند. به عنوان مثال، پاسخ‌های فیزیولوژیکی و رفلکسی مثل درد و حالت تهوع، معیار دقیق و قابل مشاهده‌ای برای ارزیابی ندارند و برای درک وضعیت، باید کاملاً به توضیحات بیمار تکیه کرد (۲۰). برای محاسبه MCID از این طریق، تغییرات نمرات بیماران در شاخص‌های ارزیابی مثل پرسشنامه‌های ناتوانی و درد به عنوان نمرات PRO و با شاخص جهانی تغییر (GRC) مقایسه می‌شود. GRC با رتبه‌بندی بیماران از +۷ (کاملاً بهتر) تا -۷ (کاملاً بدتر) به نحوی سریع، انعطاف پذیر و کارآمد اطلاعات وضعیت بیمار را بدست می‌آورد که توسط پیشکا و همکاران و براساس مقیاس هفت امتیازی لیکرت ابداع شد (۶). توجه داشته

<sup>16</sup> Patient-reported outcome

که شدت مشکلش در گذشته تا چه اندازه بوده است و تحت تأثیر درمانی که گرفته و کاهش علائم بیماری اعلام کند که: "درمان بسیار خوب و موفقیت آمیز بوده است" که در این صورت این پاسخ دارای سوگیری است (۱۵). لذا طول زمان یادآوری نباید افزایش یابد، زیرا بر نمرات GRC تأثیر منفی می‌گذارد. در برخی مطالعات رتبه‌بندی عددی شاخص لنگر از ۵ تا ۱۱، بسته به بیماری، مشکل و شاخص‌های ارزیابی لنگر مشاهده شده است (۲۲، ۲۳)؛ اما به طور کلی، به نظر می‌رسد مقیاس‌هایی با ۷ و ۱۱ امتیاز بهترین همبستگی را بین ترجیح بیمار، توانایی تشخیص کافی و قابلیت اطمینان آزمون مجدد ارائه می‌دهند (۲۱).

### نحوه محاسبه MCID بر اساس روش مبتنی بر لنگر

کوپای و همکاران (۵)، رایت و همکاران (۱۱) و موولی و همکاران (۲۴) در مطالعات خود نحوه محاسبه MCID مبتنی بر لنگر را به دو شکل مقطعی و طولی بیان کرده‌اند و باتوجه به اینکه هدف ما در این مطالعه بررسی اثربخشی مداخلات در مطالعات کارآزمایی بالینی می‌باشد در ادامه تنها به بررسی روش طولی و توضیح مختصر سه رویکرد آن می‌پردازیم.

### تغییر نمرات درون گروهی بیماران<sup>۱۷</sup>

این بیماران براساس پاسخ‌های خود به مقیاس ارزیابی جهانی یا GRC دسته‌بندی می‌شوند. سپس میانگین تغییرات نمرات  $PRO (\Delta \bar{x})$  (انتهای جدول ۱) بیمارانی که حتی تفاوت کوچکی (امتیازهای ۱، ۲، ۳) در وضعیت خود احساس کرده‌اند را به عنوان MCID انتخاب می‌کنند (۶، ۲۵). در مطالعه‌ای دیگر، امتیاز ۱ را معادل صفر در نظر گرفته و فقط میانگین تغییر نمره PRO بیمارانی که امتیاز  $\pm ۲$  و ۳ را کسب کرده‌اند، جهت تعیین MCID محاسبه می‌گردد (۲۶). به عنوان مثال، در یک مطالعه بالینی، بیماران پس از پایان مداخله درمانی پرسشنامه ناتوانی را تکمیل می‌کنند و سپس به سوال GRC که بین +۷ (کاملاً بهتر) و -۷ (کاملاً بدتر) متغیر است، پاسخ می‌دهند. در نهایت، میانگین تغییر نمرات PRO (به طور مثال امتیاز پرسشنامه ناتوانی فرد) بیمارانی که امتیاز GRC آنها  $\pm ۱$ ، ۲، ۳ است را محاسبه و به عنوان نمره MCID اعلام می‌کنند. گاهی پیش می‌آید که امتیاز پرسشنامه ناتوانی (یا

۱- خیلی خیلی جزئی بدتر (در حد بدون تغییر)	
۲- کمی بدتر	
۳- تا حدودی بدتر	
۴- به طور متوسط بدتر	
۵- کمی بیشتر از متوسط بدتر	
۶- خیلی بدتر	
۷- به طور کامل بدتر	

جدول ۳- شاخص GCI-C، شناخته شده به عنوان لنگر در میزان تغییرات مداخلات درمانی؛ پیشنهاد شده توسط گای و همکاران (۱۹۷۶).

سؤال: در مقایسه با زمان پذیرش چقدر تغییر را در خود احساس می‌کنید؟

تغییرات	امتیاز شاخص
خیلی خیلی بهتر	۱
خیلی بهتر	۲
کمی بهتر	۳
بدون تغییر	۴
کمی بدتر	۵
خیلی بدتر	۶
خیلی خیلی بدتر	۷

معمولاً لازم است که یک ارتباط ثابت بین معیار خارجی (لنگر) و ابزار اندازه‌گیری وضعیت بیمار وجود داشته باشد تا نتایج معتبری در مورد امتیازات PRO حاصل شود. این معیار نسبت به روش‌های دیگر سریع‌تر و ساده‌تر وضعیت بیمار را مشخص می‌کند، اما به دلیل تناقض در مطالعات، اعتبار و پایایی نامشخص آن و تأثیر «سوگیری یادآوری» بر پاسخ‌دهی طولانی‌مدت، مورد بحث و انتقاد قرار گرفته است (۲۱). سوگیری یادآوری زمانی اتفاق می‌افتد که بیمار آنچه اخیراً اتفاق افتاده را به بهترین وجه به یاد می‌آورد، اما گذشته را به خوبی به یاد نمی‌آورد، به طور مثال ممکن است بیمار به طور دقیق نتواند درک کند

<sup>17</sup> Within patient's score change

اینکه هیچ تفاوت آماری بین آنها وجود نداشته باشد افزایش می‌یابد و MCID به اندازه‌ی خطای اندازه‌گیری کوچک می‌شود (۲۴، ۲۸). در شکل ۳، تفاوت میانگین ناتوانی افرادی که امتیاز PRO آنها در نقطه‌ی بهتر بود با افرادی که امتیاز آنها در نقطه‌ی بدون تغییر بود ۹/۵۳ شد (۲۷). در این رویکرد، دو سطح مجاور مقیاس را بر اساس ابزار اندازه‌گیری مطالعه انتخاب می‌کنیم (۴، ۲۴)

### رویکرد مبتنی بر حساسیت و ویژگی<sup>۱۹</sup>

در این رویکرد نمره‌ای را به عنوان MCID انتخاب می‌کنیم که بهترین قدرت تمایز را در بین گروه‌های بیمار داشته باشد (یعنی نمره‌ای که بیشترین حساسیت و ویژگی را داشته باشد). در ارتباط با MCID، حساسیت نسبت بیمارانی است که تشخیص بهبودی آنها درست بوده است (تغییرات مثبت واقعی) و امتیازات PRO آنها بالاتر از نمره در نظر گرفته شده برای MCID است. ویژگی نیز آن نسبتی از بیماران است که هیچ بهبودی حاصل نکرده‌اند و این عدم تغییر به درستی شناخته شده است (تغییرات منفی واقعی) و امتیازات PRO آنها کمتر از مقدار آستانه MCID است (۳، ۶). به‌طور کلی در یک مطالعه بالینی چهار حالت رخ می‌دهد:

مثبت واقعی (TP<sup>۲۰</sup>): درمان تأثیرگذار بود و این تشخیص واقعی است.

مثبت کاذب (FN<sup>۲۱</sup>): درمان تأثیرگذار بود؛ اما به خطا آن را بی‌تأثیر اعلام کردند.

منفی واقعی (TN<sup>۲۲</sup>): درمان تأثیر خاصی نداشت و این تشخیص واقعی است.

منفی کاذب (FP<sup>۲۳</sup>): درمان تأثیری نداشت؛ اما آن را به خطا تأثیرگذار اعلام کردند.

با توجه به چهار حالت ذکر شده، در مطالعه حاضر حساسیت میزان تشخیص درست تأثیر درمان است که به آن نرخ مثبت واقعی (TPR) می‌گویند که مثبت واقعی و منفی کاذب را در برمی‌گیرد و با فرمول زیر محاسبه می‌شود:

$$Sensitivity(TPR) = \frac{TP}{TP + FN}$$

<sup>19</sup> Sensitivity- and specificity-based approach

<sup>20</sup> True Positive

<sup>21</sup> False Negative

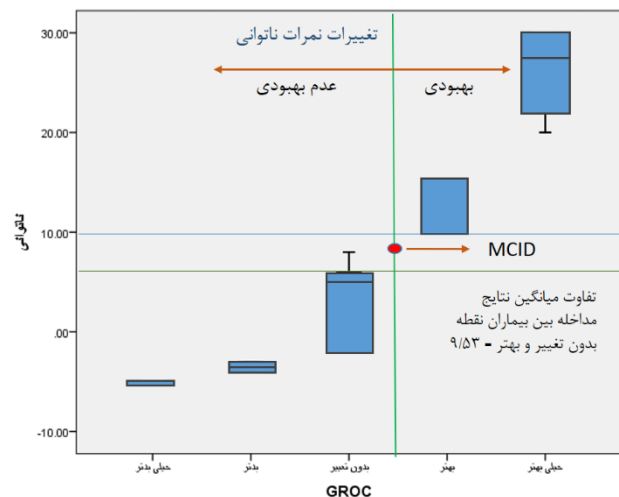
<sup>22</sup> True Negative

<sup>23</sup> False Positive

هر متغیر دیگر) بیمار نشان از بهبود دارد اما امتیازی که فرد در شاخص GRC به بهبودی می‌دهد پایین است. به طور مثال امتیاز ۱، در این حالت فارغ از نمره بالای PRO امتیاز وی محاسبه نمی‌گردد و از محاسبه میانگین خارج می‌شود.

### تغییر نمرات بین گروهی بیماران<sup>۱۸</sup>

این رویکرد براساس مقایسه‌ی تغییر نمرات PRO (مثلاً در پرسشنامه‌ی ناتوانی) در گروه‌های بیمار با پاسخ‌های مختلف به GRC است. در یک مطالعه‌ی مقطعی، MCID را برابر با تفاوت امتیاز بین دو سطح مجاور از تغییر کم و عدم تغییر در GRC در نظر گرفته‌اند (۱۹، ۲۷)؛ و در یک مطالعه‌ی طولی، بیماران وضعیت خود را با عنوان‌های «بسیار بهتر»، «بهتر»، «بدون تغییر» و «بدتر» ارزیابی کرده‌اند. در این حالت، MCID را برابر با اختلاف میانگین نمره تغییر بیماران «بهتر» و «بدون تغییر» قرار داده‌اند ( $\bar{X}_\Delta$ ). بیماران بهبودیافته  $\bar{X}_\Delta$  - بیماران بهبود نیافته. از نظر مفهومی، MCID باید نشان‌دهنده تفاوت بین دو سطح مجاور در یک مقیاس باشد و تفاوت میانگین گروهی که کمی بهبود یافته‌اند با گروهی که بهبود نیافته‌اند را نشان دهد. بنابراین، MCID به تعداد سطوح مقیاس وابسته است.



شکل ۳- تفاوت میانگین نتایج ناتوانی پس از مداخله در گروه بهبود یافته و بهبود نیافته

هرچه تعداد سطوح بیشتر باشد، اختلاف بین دو سطح مجاور کمتر می‌شود و MCID نیز کوچکتر می‌شود. اگر امتیازات دو سطح مقیاس به هم نزدیک باشند، احتمال

<sup>18</sup>Between patients score change

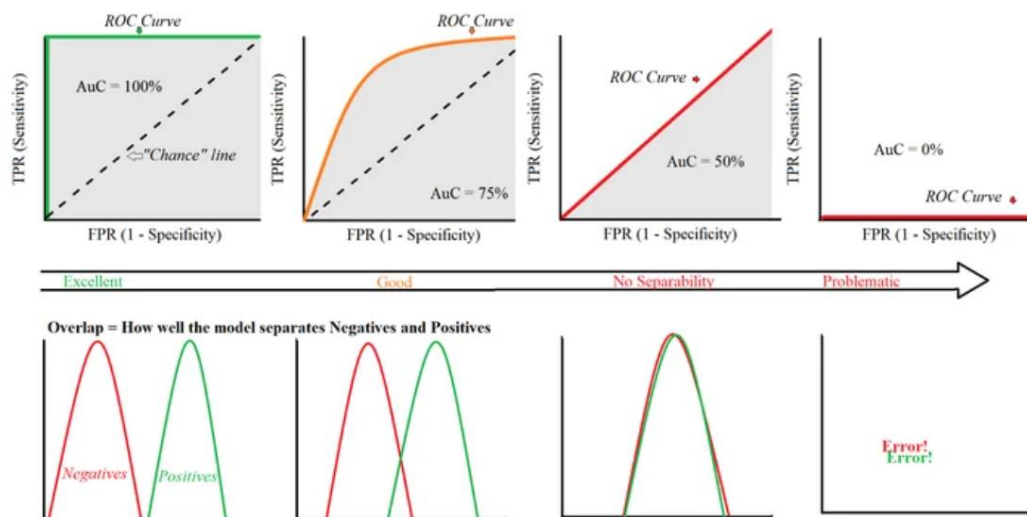
طبقه‌بندی باینری (دودویی) را دارد و از ابزارهای سنجش عملکرد یک روش آزمون به حساب می‌آید که با استفاده از آن می‌توان مفاهیمی مانند نقطه برش، حساسیت و ویژگی آن آزمون را مورد بررسی قرار داد (۷، ۲۴). منحنی راک نمودار مقادیر حساسیت (TPR) در مقابل ویژگی-۱ (FPR) است، زیرا مقدار نقطه برش از ۰ به ۱ حرکت می‌کند. در منحنی ROC، مدلی که حساسیت بالا و ویژگی بالا داشته باشد، به گوشه سمت چپ بالای طرح نزدیک می‌شود و مدلی که حساسیت و ویژگی آن کم است دارای منحنی نزدیک به خط نیم‌ساز خواهد بود (۲۷). در شکل ۳ به خوبی مدل‌های منحنی ROC به نمایش گذاشته شده است. از راست به چپ، منحنی بهترین تأثیر (در مداخله) یا تشخیص (در سنجش تست) را نشان می‌دهد و فضای زیرمنحنی آن ۱۰۰ درصد است. هر چه به سمت چپ پیکان نزدیک می‌شویم، منحنی به خط نیم‌ساز نزدیکتر و در نهایت از آن عبور می‌کند و فضای زیر منحنی کوچک می‌شود. در این حالت در مرکز نمودار مثبت کاذب یا صفر را داریم و در یک گوشه بالای نمودار مثبت واقعی (خط سبز) و در گوشه پایین نمودار منفی واقعی (خط قرمز) را داریم.

همچنین ویژگی میزان تشخیص درست و واقعی عدم درمان است که به آن نرخ منفی صحیح (TNR) می‌گویند که حالت‌های منفی صحیح و مثبت کاذب را دربر می‌گیرد و فرمول آن به شکل زیر است:

$$\text{Specificity (TNR)} = \frac{TN}{TN + FP}$$

نمره حساسیت ۱ نشان دهنده این است که همه موارد مثبت واقعی، به درستی شناسایی شده‌اند و نمره ویژگی ۱ نشان می‌دهد که همه موارد منفی واقعی کاملاً شناسایی شده‌اند. تا به امروز حساسیت یا ویژگی مطلوب و مشخص شده‌ای برای MCID تعیین نشده است و محققان حساسیت و ویژگی یکسانی را برای تعیین MCID در این روش انتخاب می‌کنند. در این روش به مانند سایر روش‌ها خطای ناشی از گزارش اشتباه بهبودی توسط بیماران وجود دارد، که همان منفی کاذب و مثبت کاذب می‌باشد (۳).

در رویکرد مبتنی بر حساسیت و ویژگی، از منحنی ROC برای شناسایی امتیاز PRO با حساسیت و ویژگی مساوی برای تمایز بین بیماران دارای «بهبودی» و «عدم بهبودی» استفاده می‌شود. منحنی ROC که از آن به عنوان «منحنی مشخصه عملکرد» نام برده می‌شود، یک نمودار گرافیکی است که توانایی بررسی و ارزیابی عملکرد سیستم



شکل ۴- انواع مدل‌های منحنی ROC

بهبودی در همه بیماران است و البته مساحت ۰/۸ تا ۰/۹ خوب و مساحت ۰/۷ تا ۰/۸ نسبتاً خوب در نظر گرفته می‌شود (مساحت پایین‌تر از ۰/۷ به عنوان تأثیر ضعیف و غیرقابل قبول شناخته می‌شوند). AUC نشان‌دهنده تغییر

دامنه احتمالات محدوده‌ی زیرمنحنی (AUC) از ۰/۵ تا ۱ متغیر است. ۰/۵ نشان‌دهنده توانایی تشخیص به دلیل شانس یا تصادف و ۱ نشان‌دهنده توانایی تشخیص درست

گاهی روش‌های آماری برای حذف اثر نمرات پیش‌آزمون مورد استفاده قرار می‌گیرد. این کنترل‌های آماری فرض می‌کنند که امتیازات پیش‌آزمون بالا به دلیل شانس یا خطا است (مثل آزمون کواریانس که در آن نمرات پیش‌آزمون به عنوان عامل کوریت کنترل می‌شوند). پس بهتر است برای اختلاف پیش‌آزمون و پس‌آزمون به جای نمرات تغییر خام از درصد تغییرات استفاده شود. استفاده از امتیازهای درصد تغییرات، نمرات پیش‌آزمون بالا را در زمانی که احتمال دارد تغییرات خیلی بزرگ دیده شوند اصلاح می‌کند. این مورد برای مقیاس‌های درد و شاخص ناتوانی مثل آسوستری و یا شاخص‌ها و ابزارهایی است که در آن تفاوت میانگین بالا بین پیش‌آزمون و پس‌آزمون نشان‌دهنده بهبودی و تفاوت میانگین پایین (کوچک) نشان‌دهنده کاهش یا عدم بهبودی است. (۶).

### بحث و نتیجه‌گیری

در مطالعه حاضر، نویسندگان به تبیین مفاهیم حداقل تفاوت مهم کلینیکالی یا MCID و مهم‌ترین رویکردهای مورد استفاده در محاسبه آن پرداختند. با بررسی متون و مطالعات مختلف در رابطه با MCID محققین مطالعه حاضر دو رویکرد مهم در این موضوع را مورد بررسی قرار دادند: رویکرد مبتنی بر لنگر و رویکرد مبتنی بر توزیع. بهترین کاربرد روش اول در زمانی است که بیمار یا افراد مورد مطالعه در دسترس باشند و دیدگاه بیمار در این روش تأثیر مستقیم دارد. کاربرد روش دوم در زمانی است که بیماران به هر دلیلی از دسترس خارج هستند و محقق یا پزشک برای برآورد اثربخشی مداخله مورد مطالعه از این روش که وابسته به ویژگی‌های آماری و مختص نمونه است، استفاده می‌کند.

به‌طور کلی، MCID به‌عنوان استاندارد نسبتاً جدید برای تعیین اثربخشی یک درمان معین و توصیف رضایت بیمار نسبت به آن درمان ارائه شده است. این روش توسط پزشکان برای تعیین پاسخ بیمار به درمان و هدایت تصمیم‌گیری بالینی و همچنین به‌عنوان یک آستانه برای تشخیص تغییرات فراتر از خطاهای تصادفی در طول درمان استفاده می‌شود. اگرچه محققان بالینی نیز از این نمره جهت تعیین آستانه درمانی برای اثربخشی مداخلات مورد بررسی استفاده می‌کنند (۳، ۴). مفیدترین مفهوم MCID این است که پزشک و بیمار متوجه این نکته شوند که آیا

یا عدم تغییر در بیماران است که منفی کاذب در مرکز و منفی واقعی در گوشه نمودار می‌باشد (۲۸). در یک مثال برای تعیین MCID بیماران مبتلا به کمردرد که یک مداخله درمانی جدید برای کاهش درد و ناتوانی دریافت کرده‌اند؛ در مرحله اول، امتیاز PRO افراد را قبل و بعد از مداخله اندازه‌گیری می‌کنیم (مانند پرسش‌نامه‌ی آسوستری) و نمرات تغییر خام یا درصد تغییرات آن‌ها را بدست می‌آوریم (تفاوت پیش‌آزمون و پس‌آزمون). سپس با کمک یک لنگر خارجی (مانند CGI-C) توزیع تغییرات بیماران را (از ۱ تا ۷) مشخص می‌کنیم. سپس با توجه به سیستم دودویی منحنی ROC افراد را به دو گروه بهبود یافته و عدم بهبودی تقسیم می‌کنیم. افرادی که در شاخص CGI-C، گزینه‌های ۱-۳ را انتخاب کردند افراد بهبود یافته (مثبت) و افرادی که گزینه‌های ۴-۷ را انتخاب کردند جز دسته‌ی عدم بهبودی (منفی) محسوب می‌شوند. سپس نمودار ROC را با استفاده از یکی از نرم افزارهای مختلف آماری مثل SPSS، STATA، R و ... محاسبه می‌کنیم. در جدول مختصات نیز چندین نقطه برش داده شده است که برای هر کدام مقدار حساسیت و مقدار (ویژگی - ۱) آورده شده است. برای تعیین بهینه‌ترین نقطه برش از شاخص یو دِن استفاده می‌کنیم. بر مبنای این شاخص برای انتخاب یک نقطه برش بهینه باید مقداری را مشخص کنیم که بالاترین حساسیت و پایین‌ترین مقدار را برای (ویژگی - ۱) داشته باشد، معمولاً این نقطه را به عنوان MCID انتخاب می‌کنند (۱۹، ۲۸).

برای درک امتیاز MCID از مفهوم «روش تغییر میانگین» استفاده می‌کنیم. بر اساس این روش، اگر تفاوت مطلق بین دو میانگین اختلاف نمره قبل و بعد دو گروه بهبود یافته و بهبود نیافته یک عدد مثبت باشد، MCID برای بهبودی را نشان می‌دهد (به عنوان مثال، تسکین درد) و اگر منفی باشد، منعکس‌کننده MCID برای بدتر شدن است. MCID هایی که به این ترتیب تعیین می‌شوند برای کاربرد در گروه معتبر هستند اما در سطح فردی معتبر نیستند (۱، ۲۹).

### وابستگی به نمرات پیش‌آزمون و کاهش آن

تغییرات PRO به نمرات پیش‌آزمون وابسته است یعنی بیماران با سطح ناتوانی بیشتر در پیش‌آزمون، بهبودی بیشتری را نشان می‌دهند (۳۰، ۳۱). برای حل این مسئله،



(۵، ۱۱).

نقاط ضعف روش‌های مبتنی بر لنگر این است که خطای اندازه‌گیری را در نظر نمی‌گیرد و اگر نمرات خیلی بهم نزدیک باشند نقطه برش وارد محدوده ۵ درصد خطای اندازه‌گیری می‌شود؛ و همچنین ممکن است در این روش سوگیری یادآوری به وجود بیاید. اما نقطه قوت این معیار این است که نسبت به روش‌های دیگر خیلی سریع‌تر و ساده‌تر وضعیت بیمار را مشخص می‌کند و با حفظ شرایط صحیح مطالعه می‌توان نتایج تقریباً مشخص‌تری از آن بدست آورد، زیرا دیدگاه بیمار در این روش در نظر گرفته می‌شود (۵، ۱۱، ۱۴). درک مفاهیم مذکور در مطالعه حاضر، محاسبه و کاربرد نمره MCID را تسهیل می‌کند. محققین باید عوامل مؤثر بر نمرات MCID را که مختص جمعیت مورد مطالعه است به خوبی درک کنند. استفاده از نمره MCID بهتر است تا زمانی که بتوان به یک اجماع در محاسبه برای رفع محدودیت‌های روش‌شناسی، جمعیت و وابستگی به پیش‌آزمون دست یافت محتاطانه باشد.

پیشنهاد پژوهش حاضر برای محققین آینده این است که تعیین یک ابزار ایده‌آل جهت مشخص نمودن حداقل تفاوت مهم بالینی مرتبط با یک مداخله معین را، با در نظر گرفتن محدودیت‌ها و کاستی‌های روش‌های فعلی، مورد بررسی قرار دهند. همچنین نیاز به رفع مشکلات روش‌های تعیین نمره MCID مثل تفاوت نمره MCID در روش‌های آن، از دست دادن دیدگاه بیمار، رابطه بین تغییرات نمرات قبل و بعد از درمان، و همچنین در نظر گرفتن هزینه درمان برای بیمار هنگام تعریف MCID، مشهود است؛ و در آخر پس از بررسی مطالعات مختلف در این حوزه، به نظر می‌رسد که مطالعات آینده باید تحقیقات بیشتری در جهت ایجاد مقادیر MCID برای ابزارهای خاص PRO که معمولاً در تحقیقات مشکلات مرتبط با مفاصل مخصوصاً ستون فقرات و کمردرد و دیگر شرایط بالینی (مانند مقیاس‌های درد، شاخص ناتوانی Oswestry، و فرم کوتاه مطالعه نتایج پزشکی و...) استفاده می‌شوند، انجام دهند.

### تشکر و قدردانی

این مطالعه حاصل مرور و بررسی مطالعات مرتبط با روش‌شناسی و مفاهیم رویکرد MCID و روش‌های مرتبط با آن است. از همه کسانی که در این مطالعه یاری رساندند تشکر و قدردانی می‌نمایم.

مداخله مؤثر بوده است یا خیر؟ معمولاً انتخاب روش مورد استفاده برای محاسبه MCID می‌تواند به طور قابل توجهی بر مقادیر گزارش شده MCID تأثیر بگذارد. پزشکان و محققان بالینی باید دیدگاه بیمار از تغییر را هنگام تعیین نمره MCID در نظر بگیرند، زیرا این نمره به عنوان حداقل مقدار تغییری که برای بیمار مهم است تعریف می‌شود. برآوردهای MCID بهتر است که بر اساس دیدگاه بیمار و با استفاده از روش‌های مبتنی بر لنگر باشد، زیرا روش‌های مبتنی بر توزیع، دیدگاه بیمار از اهمیت بالینی را در نظر نمی‌گیرند، اما گاهی در شرایطی که محقق امکان استفاده از این روش را به هر دلیلی نداشته باشد، روش‌های مبتنی بر توزیع مورد استفاده قرار می‌گیرد (۹-۵).

اگرچه توسعه MCID توانایی محققین برای تعیین روش‌های مراقبتی و درمانی که نتایج بهتری را برای جمعیت‌های همگن ارائه می‌کند، بهبود بخشیده است و همچنین روش‌های زیادی برای محاسبه‌ی آن وجود دارد، اما گستردگی آن گاهی منجر به وجود مشکلاتی در تفسیر و کاربرد پیامدهای مطالعاتی می‌شوند (۱)، به طور کلی، هر کدام از این روش‌ها نقاط قوت و ضعفی دارند؛ به عنوان مثال تعاریف MCID هزینه درمان برای بیمار را در نظر نمی‌گیرد و یا تغییر در امتیازات PRO یک ارزش ثابت و مشخص جهانی ندارد و به وضعیت نمرات پیش‌آزمون بیمار بستگی دارد؛ در نتیجه، هیچ روش پذیرفته شده جهانی برای تعیین MCID وجود ندارد و همین موضوع گاهی منجر به ایجاد مقادیر مختلفی از نمره MCID گزارش شده برای اندازه‌گیری یک پیامد می‌شود، به طور مثال نمره MCID کم‌درد در روش A ممکن است ۲ باشد اما در روش B می‌شود ۲/۸۹؛ لذا نمرات MCID در روش‌های مختلف می‌توانند تا حدودی به هم نزدیک باشند اما هیچ‌گاه به یک امتیاز ختم نمی‌شوند (۶-۴). همچنین به طور اختصاصی هر کدام از رویکردهای تعیین MCID دارای کاستی‌های مختص به خود است، به طور مثال، نقاط ضعف روش‌های مبتنی بر توزیع این است که به نظرسنجی از بیماران در رابطه با تغییرات به وجود آمده نمی‌پردازد و دیگری نیز مرتبط با تأثیر حجم نمونه بر آن است. همچنین نقاط قوت این روش این است که توانایی بالایی در محاسبه تغییرات به وجود آمده فراتر از سطوح تغییرات تصادفی دارد و در زمانی که بیماران در دسترس نیستند کارایی دارد

## تضاد منافع

ما نویسندگان مطالعه حاضر اظهار می‌داریم که هیچ‌گونه تضاد منفعی وجود ندارد.

## References:

- Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *Jama*. 1994;272(2):125-8.
- Roach KE. Measurement of health outcomes: reliability, validity and responsiveness. *JPO: Journal of Prosthetics and Orthotics*. 2006;18(6):P8-P12.
- Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Controlled clinical trials*. 1989;10(4):407-15.
- Cook CE. Clinimetrics corner: the minimal clinically important change score (MCID): a necessary pretense. *Journal of Manual & Manipulative Therapy*. 2008;16(4):82E-3E.
- Copay AG, Subach BR, Glassman SD, Polly Jr DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *The Spine Journal*. 2007;7(5):541-6.
- Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, et al. Minimal clinically important differences: review of methods. *The Journal of rheumatology*. 2001;28(2):406-12.
- Lassere M, van der HEIJDE D, Johnson KR. Foundations of the minimal clinically important difference for imaging. *The Journal of rheumatology*. 2001;28(4):890-1.
- Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *Journal of clinical epidemiology*. 2003;56(5):395-407.
- Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Physical therapy*. 2006;86(5):735-43.
- Beaton DE. Understanding the relevance of measured change through studies of responsiveness. *Spine*. 2000;25(24):3192-9.
- Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). *Journal of Manual & Manipulative Therapy*. 2012;20(3):160-6.
- Wu C-y, Chuang L-l, Lin K-c, Hong W-h. Responsiveness, minimal detectable change, and minimal clinically important difference of the Nottingham Extended Activities of Daily Living Scale in patients with improved performance after stroke rehabilitation. *Archives of physical medicine and rehabilitation*. 2011;92(8):1281-7.
- Cohen J. *Statistical power analysis for the behavioral sciences*: Academic press; 2013.
- Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? *Journal of biopharmaceutical statistics*. 2004;14(1):97-110.
- Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical care*. 2003;582-92.
- Bolton JE. Sensitivity and specificity of outcome measures in patients with neck pain: detecting clinically significant improvement. *Spine*. 2004;29(21):2410-7.
- Guyatt G, Osoba D, Wu A, Wyrwich K, Norman G. Clinical Significance Consensus Meeting G. Methods to explain the clinical significance of health status measures *Mayo Clin Proc*. 2002;77(4):371-83.
- Yalfani A, Bigdeli N, Gandomi F. Comparing the effects of suspension and isometric-isotonic training on postural stability, lumbopelvic control, and proprioception in women with diastasis recti abdominis: a randomized, single-blinded, controlled trial. *Physiotherapy Theory and Practice*. 2022:1-13.
- Hägg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *European Spine Journal*. 2003;12:12-20.
- Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. 1992.
- Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *Journal of Manual & Manipulative Therapy*. 2009;17(3):163-70.
- Stewart M, Maher CG, Refshauge KM, Bogduk N, Nicholas M. Responsiveness of pain and disability measures for chronic whiplash. *Spine*. 2007;32(5):580-5.
- Haspelslagh SR, Van Suijlekom HA, Lamé IE, Kessels AG, van Kleef M, Weber WE. Randomised controlled trial of cervical radiofrequency lesions as a treatment for cervicogenic headache [ISRCTN07444684]. *BMC anesthesiology*. 2006;6:1-11.
- Mouelhi Y, Jouve E, Castelli C, Gentile S. How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. *Health and quality of life outcomes*. 2020;18:1-17.
- Sedaghat AR. Understanding the minimal clinically important difference (MCID) of patient-reported outcome measures. *Otolaryngology-Head and Neck Surgery*. 2019;161(4):551-60.

26. JUNIPER E. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol*. 1994;47:81-7.
27. Kulkarni AV. Distribution-based and anchor-based approaches provided different interpretability estimates for the Hydrocephalus Outcome Questionnaire. *Journal of clinical epidemiology*. 2006;59(2):176-84.
28. Chuang L-H, Garratt A, Brealey S. Comparative responsiveness and minimal change of the Knee Quality of Life 26-item (KQoL-26) questionnaire. *Quality of life research*. 2013;22:2461-75.
29. Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *Journal of clinical epidemiology*. 2017;82:128-36.
30. Riddle DL, Stratford PW, Binkley JM. Sensitivity to change of the Roland-Morris back pain questionnaire: part 2. Physical therapy. 1998;78(11):1197-207.
31. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris back pain questionnaire: part 1. Physical therapy. 1998;78(11):1186-96.