

MASTERCLASS

DiTA: helping you search for evidence of diagnostic test accuracy in physical therapy



Mark A. Kaizik^{a,*}, Mark J. Hancock^b, Robert D. Herbert^{a,c}

^a School of Biomedical Sciences, Faculty of Medicine and Health, University of New South Wales, Sydney, Australia

^b Faculty of Medicine and Health Sciences, Macquarie University, Sydney, Australia

^c Neuroscience Research Australia (NeuRA), Sydney, Australia

Received 19 March 2024; received in revised form 30 April 2024; accepted 8 May 2024

Available online 20 May 2024

KEYWORDS

Accuracy;
Bibliographic
database;
Diagnosis;
Physical therapy

Abstract

Background: Physical therapists use diagnostic tests in a variety of settings. Choosing the best diagnostic test to apply in a particular situation can be difficult. The choice of diagnostic test should be informed, at least in part, by evidence of test accuracy. Finding evidence of diagnostic test accuracy has, until recently, been challenging. Ideally, there would exist a database that comprehensively indexes evidence on diagnostic tests relevant to physical therapy practice, is free to access, and is easy to use.

Objective: This Masterclass will describe the DiTA (*Diagnostic Test Accuracy*) database (dita.org.au) including its development and search interface, and provide advice on how to search and retrieve records.

Discussion: DiTA indexes more than 2400 primary studies and systematic reviews of diagnostic test accuracy relevant to physical therapy practice. Users can search DiTA using text fields and dropdown lists to find evidence of diagnostic test accuracy. The database is freely accessible on the internet. Since its launch, DiTA has been accessed from almost every country in the world, the largest number of searches having been conducted from Brazil.

© 2024 The Authors. Published by Elsevier España, S.L.U. on behalf of Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Diagnostic tests are used to assess the presence or absence of specific pathologies.¹ It is helpful to distinguish diagnostic tests from other types of (non-diagnostic) tests used in

clinical practice. Non-diagnostic tests are used for purposes such as assessing function, assessing impairment or monitoring progress (e.g., range of motion tests, muscle strength tests); or predicting future symptoms, conditions, or response to treatment (e.g., the STarT Back Tool for low back pain²). This Masterclass paper focuses on diagnostic tests. It describes the DiTA database and explains how DiTA can help physical therapists find evidence of diagnostic test accuracy.

* Corresponding author at: Level 7, 4 Martin Place, Sydney, NSW 2000, Australia.

E-mail: m.kaizik@student.unsw.edu.au (M.A. Kaizik).

<https://doi.org/10.1016/j.bjpt.2024.101082>

1413-3555/© 2024 The Authors. Published by Elsevier España, S.L.U. on behalf of Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Diagnostic tests are used in most subdisciplines of physical therapy.³ An example from musculoskeletal physical therapy is the use of Lachman's test for rupture of the anterior cruciate ligament of the knee⁴; an example from cardio-respiratory physical therapy is the PUMA COPD diagnostic questionnaire⁵; and an example from women's health physical therapy is the pad test for urinary incontinence.⁶

Clinicians conduct diagnostic tests in the expectation that the findings of the diagnostic test will increase certainty about whether a particular pathology is present or absent.⁷ A physical therapist often has multiple choices when selecting a diagnostic test to apply. However, the process of choosing between tests can be bewildering and it is often not clear which test is the best to choose.⁸ An example arises when a physical therapist assesses for the presence or absence of sacroiliac joint pathology in patients with low back pain.⁹ The physical therapist could consider using the distraction test, the compression test, the thigh thrust test¹⁰⁻¹²; Gaenslen's test^{11,12}; the sacral thrust test¹³; the patient's report of pain over the area of the sacroiliac joint^{14,15}; and various composites of these and other tests.^{11,13,14,16} A clinician's choice of a diagnostic test to apply can be influenced by many factors such as pattern recognition and heuristics,¹⁷ the patient's preference for a particular test,¹⁸ availability or ease of access to the test,¹⁹ and fear of litigation.²⁰ Ideally the choice of test should be influenced by evidence of diagnostic test accuracy.²¹

Evidence of diagnostic test accuracy

Evidence of diagnostic test accuracy can be found in reports of primary studies of diagnostic test accuracy and systematic reviews of primary studies of diagnostic test accuracy. In a typical primary study, both an index test and a reference standard are applied to subjects who are suspected of having a particular pathology and these test results are compared. The index test is the diagnostic test whose accuracy is being evaluated. The ideal reference standard (sometimes called a "gold standard") is a diagnostic test that perfectly classifies the subject as either having or not having the pathology. In practice, it is rare to find perfect or near-perfect reference standards, so the reference standard used is generally the best available diagnostic test for the pathology being assessed.²² Researchers conduct studies to determine the accuracy of index tests because, while index tests are usually expected to be less accurate than the reference test, they may be preferred over reference tests, in some clinical contexts, because they are less expensive, more accessible, easier to conduct, less invasive, safer, or less painful than the reference test. For example, Lachman's test has been used as an index test for the diagnosis of anterior cruciate ligament rupture, rather than the best reference test of arthroscopy. However, if an index test is to be used in clinical practice it must be sufficiently accurate to aid, rather than confuse, diagnosis.

Primary studies of diagnostic test accuracy compare the findings of the index test to the reference standard. The similarity of those findings provides a measure of the diagnostic accuracy of the index test. Accuracy can be quantified and reported in a variety of ways. Most commonly accuracy is reported in terms of sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood

ratio, and negative likelihood ratio.²³ Other accuracy measurements include Youden's index,²⁴ diagnostic odds ratios,²⁵ receiver operating characteristic (ROC) curves and the area under ROC (AUROC) curves,²⁶ and a measure more simply known as accuracy, which is the ratio of the combined true positive and true negative cases to the total number of cases evaluated.²⁷

When more than one primary study has investigated the accuracy of a particular index test, researchers may conduct a review of those studies. Ideally the review uses a systematic review methodology. Systematic reviews of primary studies of diagnostic test accuracy use similar methods to systematic reviews of primary studies of intervention. Ideally, the review protocol is pre-specified, a research question is clearly posed, a comprehensive search for primary studies is conducted to retrieve and select appropriate studies, and the quality of the studies is assessed. Results are then summarised and conclusions are drawn.²⁸⁻³⁰ In systematic reviews of diagnostic test accuracy studies, data from multiple primary studies are pooled with the aim of providing more precise estimates of test accuracy such as sensitivity or specificity.^{7,31,32}

The alternative to a systematic review is a narrative review. Narrative reviews also summarise the evidence provided by primary research studies but typically fail to specify the sources and search strategy of included studies, do not explicitly nominate the criteria used to select studies for inclusion in the review, and summarise data qualitatively rather than quantitatively.³³ Consequently, it is thought by many methodologists that narrative reviews are more likely to generate biased conclusions than systematic reviews. For that reason, systematic reviews are generally preferred to narrative reviews.

The DiTA database

Ideally, there would exist a database that provides a comprehensive listing of physical therapy-related primary diagnostic test accuracy studies and related systematic reviews, is freely accessible to all, and is easy to use. In the past, it has been challenging for physical therapists to easily access evidence of diagnostic test accuracy. Thus, while PubMed is a comprehensive and freely available database of biomedical literature, it is not easy to find evidence on diagnostic test accuracy on PubMed because it indexes over 36 million records of which only a very small proportion are studies of diagnostic test accuracy.³⁴ The Cochrane Collaboration publishes a freely available register of systematic reviews, some of which are diagnostic test accuracy reviews, but it contains few reviews of diagnostic test accuracy relevant to physical therapy.³⁵ The Physiotherapy Evidence Database (PEDro; pedro.org.au) is a database that indexes primary studies and systematic reviews relevant to physical therapy; however, PEDro indexes evidence of the effects of interventions and not on the accuracy of diagnostic tests.³⁶

In response to the need for a comprehensive, free, easy to use database of studies of diagnostic test accuracy relevant to physical therapy, the DiTA database³⁷ (Diagnostic Test Accuracy database; dita.org.au) was launched in 2019. It was developed over a period of four and a half years by a team that was largely made up of Steering Committee members and volunteers of PEDro. DiTA's architecture, website,

SEARCH

DiTA, Diagnostic Test Accuracy, is a free database of primary studies and systematic reviews evaluating diagnostic tests used by physiotherapists. You can search the database for bibliographic details and links to full-text using this search page. DiTA is produced by the [PEDro Partnership \(Institute for Musculoskeletal Health, School of Public Health at The University of Sydney\)](#) and is hosted by [Neuroscience Research Australia \(NeuRA\)](#).

[Start new search](#) [Search help](#)

Abstract & Title

Title Only

Subdiscipline

Body part

Pathology

Type of index test

Name of index test

Type of reference test

Name of reference test

Method

Author

Source

Published since (YYYY)

New records added since (DD/MM/YYYY)

Return

records at a time

When searching:

Match all search terms (AND) Match any search term (OR)

[Start search](#)

Fig. 1 DiTA search page.

search interface, search fields, and layout are similar to the PEDro database, so regular users of PEDro will find DiTA easy to use. Following a large scale initial search of the literature to seed the database,³ DiTA's records have been updated on a monthly basis since July 2019 by staff with specific expertise in searching the physical therapy research literature to identify relevant studies using DiTA's published inclusion criteria.³⁸ For a paper to be indexed on DiTA, it must evaluate a diagnostic test procedure relevant to physical therapy. The diagnostic test procedure must be able to be performed by a physical therapist and not just produce results that a physical therapist would use (so, for example, a study evaluating the accuracy of a radiograph would not be included). There are no language restrictions for the literature indexed on DiTA.

To search the database, users must first navigate to the freely available DiTA search page (search.dita.org.au; Fig. 1). On this page, the user can nominate a search query by entering text into search fields or selecting dropdown list items. Fields include "Abstract & Title" to search for search terms in the abstract or title of indexed papers; "Subdiscipline" to search for papers related to a specific subdiscipline in physical therapy such as 'neurology' or 'cardiothoracics'; "Body part" to search for papers related to a specific part of the body such as 'head or neck', or 'perineum or genito-urinary system'; and "Pathology" to search for papers related to broad categories of pathology such as 'muscular' or 'articular'. Characteristics of the index test and reference standard can also be searched. For example, "Type of index test" allows the user to search for a specific type of diagnostic test being evaluated in a study such as a 'questionnaire', a 'physical examination' procedure, a 'health technology' like ultrasound or spirometry, or an index test that is a mixture of any of these types ('mixed'). Users can also specify what type ("Method") of indexed paper they want their search to query ('primary study' or 'systematic review'), as well as search for a specific author name ("Author") or a journal name ("Source") that may help them retrieve publications of interest.

Users have an option at the bottom of the search page to combine all their search field queries using the Boolean operators of 'AND' or 'OR'. 'Match all search terms (AND)' is the default option. If more than one word is entered into a text field, the words in that specific field are automatically combined using the Boolean operator 'AND' even if the user has also selected to combine all their search field queries at the bottom of the search page using 'OR'. When DiTA automatically combines search terms in a specific field using 'AND' this way, each word must be present in that queried field (e.g. the "Title" field) of a record for it to be retrieved although the words don't need to be adjacent to each other. Alternatively, inclusion of more than one search term within inverted commas (" ") will tell DiTA to only look for records containing all of those terms in that order. Terms can be entered into a text field using wildcards (e.g. "**") that search for variations of words in that field. Searching for "edema" in the "Title" field will retrieve records that have titles containing 'edema', 'oedema', 'lymphedema' or 'lymphoedema'.

If records exist on the database that match the search query, they will be listed in table format showing the "Title" and the "Method" of research used for each displayed record ('primary study' or 'systematic review'). Specific records can be selected from this list to view in more detail at a later

time. Alternatively, more details of a record can be viewed directly from the table by clicking on the record's title. The extra details include the authors, source, and abstract of the paper (if copyright release has been granted). DiTA does not give direct access to whole papers. However, DiTA does provide links to online platforms that may (or may not) provide the full electronic manuscript. Some papers are available in full text for free (Table 1).

DiTA has a few useful ancillary features. Search results can be saved and exported by email. The emailed search results can be imported into reference management tools like Endnote. DiTA also allows the user to set up automatic e-mail notification of newly indexed evidence. The DiTA website provides two online tutorials ("Is this study valid?" and "How can I use evidence of diagnostic text accuracy?") to help users interpret the evidence they find.

The authors recently conducted a user experience study of the DiTA website and search interface (manuscript under review). That study found participants rated DiTA above average compared to similar web platforms. Participants also reported that they quickly learned to use DiTA. However, participants were commonly observed to use DiTA in a suboptimal way. Examples included the unnecessary use of multiple search fields when fewer would have been more effective, not using the most appropriate search field available (for example, using "Title Only" when using the "Abstract & Title" field was more appropriate), incorrect spelling of search terms, and mis-specification of DiTA-specific search syntax such as wildcards (e.g. *) and Boolean operator functions. To use DiTA most efficiently, users of DiTA should usually enter data into only one or two search fields in an initial search. If only one field is used then it is best to use the "Abstract & Title" field. When using the "Abstract & Title" field, it is best to use two or three simple search terms. The search terms should be separated with spaces – not with Boolean operators ('AND' and 'OR').

Table 1 Top 4 tips on how to search DiTA for evidence of diagnostic test accuracy.

1	Limit the initial search to using just one or two search fields. (Use the "Abstract & Title" field if only using one field to search.)
2	When using the "Abstract & Title" field, just use two or three simple search terms separated with spaces and not with Boolean operators. Boolean operator functions are best controlled using the options at the bottom of the search page.
3	Check all spelling is correct, particularly the use of apostrophes for specific index test names (e.g. search for "Lachman's" or "Lachman'", not "Lachmans").
4	If too many records are found, navigate back to the initial search query by clicking "Continue Searching" at the top of the "Search results" page, edit the initial search query by adding another search term to an unused search field, and check 'Match all search terms (AND)' is selected at the bottom of the search page.

Spelling should be checked, especially the use of apostrophes for specific index test names such as ‘Lachman’s test’ (‘Lachman’s’ retrieves 51 records whereas ‘Lachmans’ retrieves 0 records). If the initial search retrieves too many records, subsequent searches can refine the initial search by

using one or more of the unused search fields and checking that ‘AND’ is selected at the bottom of the search page to make the query more specific. Users can navigate back to the initial search query by clicking “Continue Searching” at the top of the “Search results” page (Fig. 2).

DiTA
Diagnostic Test Accuracy

SEARCH BROWSE LEARN ABOUT

Search results

Display selected records

Click on a title to view details of that record. If your search has returned too many or too few records you may need to change your search terms (use the back button in your browser or the "Continue searching" link). To display a list of records from one or a series of searches, click on Select and then Display Selected Records

Found 15 records

Title	Method	Select
Special tests for assessing meniscal tears within the knee: a systematic review and meta-analysis	systematic review	Select
Der verletzte meniskus: wie sicher ist die klinische untersuchung? Eine metaanalyse (Value of the clinical examination in suspected meniscal injuries. A meta-analysis) [German]	systematic review	Select
A meta-analysis examining clinical test utilities for assessing meniscal injury	systematic review	Select
The impact of body mass index on the accuracy of the physical examination of the knee	primary study	Select
Diagnostic value of clinical tests and MRI for meniscal injury in patients with anterior cruciate ligament injury: case series study	primary study	Select
Should we prefer magnetic resonance imaging to physical examination in meniscal tears	primary study	Select
Comparison of accuracy in expert clinical examination versus magnetic resonance imaging and arthroscopic exam in diagnosis of meniscal tear	primary study	Select
Comparison of Thessaly Test with joint line tenderness and McMurray Test in the diagnosis of meniscal tears	primary study	Select
Comparison of diagnostic accuracy of physical examination and MRI in the most common knee injuries	primary study	Select
Diagnostic accuracy of the thessaly test, standardised clinical history and other clinical examination tests (Apley's, McMurray's and joint line tenderness) for meniscal tears in comparison with magnetic resonance imaging diagnosis	primary study	Select
Validity of the Thessaly test in evaluating meniscal tears compared with arthroscopy: a diagnostic accuracy study	primary study	Select

Fig. 2 DiTA search results page.

On the PEDro database, users can view quality rating scores for each randomized controlled trial indexed on the database. The scores on PEDro are generated using the PEDro scale.³⁹ This scale gives a score out of 10 which is intended to represent the methodological quality of randomized controlled trials in physical therapy. When displaying the search results for users, PEDro orders the records for randomized controlled trials from highest to lowest score to help users find the best quality evidence amongst the retrieved studies.^{40–42} It would be convenient if the DiTA database could incorporate similar features. Many quality assessment tools have been designed for the purpose of rating the quality of primary studies of diagnostic test accuracy.^{43,44} Unfortunately, these tools have been shown to be unreliable, which implies that they should not be trusted to provide robust evidence of study quality.⁴⁵ For that reason, DiTA unlike PEDro, does not provide quality rating scores for primary studies, and search results are not returned in order of presumed methodological quality.

As of March 2024, there were 2185 primary studies of diagnostic test accuracy and 277 systematic reviews of primary studies of diagnostic test accuracy indexed on DiTA. The first primary study was published in 1951⁴⁶ and the first systematic review was published in 1995.⁴⁷ The studies have been reported in 19 different languages, with a large majority in English (2315/2462; 94%). Within the first four years of its launch, the DiTA database had been accessed from almost every country in the world. Brazil was by far the largest user of DiTA: 31% of search sessions in that period were conducted from Brazil.

Conclusion

Diagnostic tests are widely used by physical therapists. It can be difficult to choose which diagnostic tests to apply in a particular situation. The choice of diagnostic test should be informed in part by evidence of test accuracy. Until recently, this evidence was not easy to find. However, the launch of the DiTA database in 2019 has made it easier for physical therapists to find evidence of diagnostic test accuracy. DiTA indexes over 2400 primary studies and systematic reviews of diagnostic test accuracy relevant to physical therapy practice. Access to the database is free. DiTA is used widely across the world, with the largest number of searches being conducted from Brazil.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors affirm that they have no financial affiliation (including research funding) or involvement with any commercial organization that has a direct financial interest in any matter included in this manuscript, except as disclosed in an attachment and cited in the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Kaizik MA, Hancock MJ, Herbert RD. DiTA: a database of diagnostic test accuracy studies for physiotherapists. *J Physiother*. 2019;65(3):119–120.
2. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008;59(5):632–641.
3. Kaizik MA, Hancock MJ, Herbert RD. A description of the primary studies of diagnostic test accuracy indexed on the DiTA database. *Physiother Res Int*. 2020;25(4):e1871.
4. Sokal PA, Norris R, Maddox TW, Oldershaw RA. The diagnostic accuracy of clinical tests for anterior cruciate ligament tears are comparable but the Lachman test has been previously overestimated: a systematic review and meta-analysis. *Knee Surg Sports Traumatol Arthrosc*. 2022;30(10):3287–3303.
5. Lopez Varela MV, Montes de Oca M, Wehrmeister FC, Rodriguez C, Ramirez L, Menezes A. External validation of the PUMA COPD diagnostic questionnaire in a general practice sample and the PLATINO study population. *Int J Chron Obstruct Pulmon Dis*. 2019;14:1901–1911.
6. Medeiros Araujo C, de Moraes NR, Sacomori C, de Sousa Dantas D. Pad test for urinary incontinence diagnosis in adults: systematic review of diagnostic test accuracy. *NeuroUrol Urodyn*. 2022;41(3):696–709.
7. Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149(12):889–897.
8. Kosack CS, Page AL, Klatser PR. A guide to aid the selection of diagnostic tests. *Bull World Health Organ*. 2017;95(9):639–645.
9. Laslett M, Williams M. The reliability of selected pain provocation tests for sacroiliac joint pathology. *Spine (Phila Pa 1976)*. 1994;19(11):1243–1249.
10. Albert H, Godskesen M, Westergaard J. Evaluation of clinical tests used in classification procedures in pregnancy-related pelvic joint pain. *Eur Spine J*. 2000;9(2):161–166.
11. Cook C, Massa L, Harm-Ernandes I, et al. Interrater reliability and diagnostic accuracy of pelvic girdle pain classification. *J Manipulative Physiol Ther*. 2007;30(4):252–258.
12. Werner CM, Hoch A, Gautier L, König MA, Simmen HP, Osterhoff G. Distraction test of the posterior superior iliac spine (PSIS) in the diagnosis of sacroiliac joint arthropathy. *BMC Surg*. 2013;13:52.
13. Laslett M, Aprill CN, McDonald B, Young SB. Diagnosis of sacroiliac joint pain: validity of individual provocation tests and composites of tests. *Man Ther*. 2005;10(3):207–218.
14. Dreyfuss P, Michaelsen M, Pauza K, McLarty J, Bogduk N. The value of medical history and physical examination in diagnosing sacroiliac joint pain. *Spine (Phila Pa 1976)*. 1996;21(22):2594–2602.
15. Murakami E, Aizawa T, Noguchi K, Kanno H, Okuno H, Uozumi H. Diagram specific to sacroiliac joint pain site indicated by one-finger test. *J Orthop Sci*. 2008;13(6):492–497.
16. Young S, Aprill C, Laslett M. Correlation of clinical examination characteristics with three sources of chronic low back pain. *Spine J*. 2003;3(6):460–465.
17. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Am J Med*. 2008;121(5A (Suppl)):S2–23.
18. Espeland A, Baerheim A, Albrektsen G, Korsbrekke K, Larsen JL. Patients' views on importance and usefulness of plain radiography for low back pain. *Spine (Phila Pa 1976)*. 2001;26(12):1356–1363.
19. Espeland A, Baerheim A. Factors affecting general practitioners' decisions about plain radiography for back pain: implications for classification of guideline barriers – a qualitative study. *BMC Health Serv Res*. 2003;3:8.

20. Studdert DM, Mello MM, Sage WM, et al. Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. *J Am Med Assoc.* 2005;293(21):2609–2617.
21. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based medicine: How to Practice and Teach EBM.* 2nd ed. Edinburgh: Churchill Livingstone; 2000.
22. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol.* 2009;62(8):797–806.
23. Davidson M. The interpretation of diagnostic tests: a primer for physiotherapists. *Aust J Physiother.* 2002;48(3):227–232.
24. Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3(1):32–35.
25. Eusebi P. Diagnostic accuracy measures. *Cerebrovasc Dis.* 2013;36(4):267–272.
26. Bewick V, Cheek L, Ball J. Statistics review 13: receiver operating characteristic curves. *Crit Care.* 2004;8(6):508–512.
27. Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. *Emerg (Tehran).* 2015;3(2):48–49.
28. Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. *Syst Rev.* 2013;2:82.
29. Deeks JJ, Wisniewski S, Davenport C. Chapter 4: guide to the contents of a Cochrane Diagnostic Test Accuracy protocol. In: Deeks JJ, Bossuyt PM, Gatsonis C, *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0.*: The Cochrane Collaboration; 2013. Available from: <http://srdta.cochrane.org/>. Accessed 6 March 2024.
30. Devillé WL, Buntinx F, Bouter LM, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol.* 2002;2:9.
31. Reitsma JB, Moons KG, Bossuyt PM, Linnet K. Systematic reviews of studies quantifying the accuracy of diagnostic tests and markers. *Clin Chem.* 2012;58(11):1534–1545.
32. Leeflang MMG. Systematic reviews and meta-analyses of diagnostic test accuracy. *Clin Microbiol Infect.* 2014;20(2):105–113.
33. Cook DJ. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med.* 1997;126(5):376–380.
34. US National Library of Medicine. PubMed. *National Institutes of Health.* <https://pubmed.ncbi.nlm.nih.gov/>. Accessed 6.03.24.
35. Cochrane Collaboration. Cochrane Database of Systematic Reviews. <https://www.cochranelibrary.com/cdsr/reviews>. Accessed 6.03.24.
36. Physiotherapy Evidence Database (PEDro). PEDro. <http://www.pedro.org.au>. Accessed 6.03.24.
37. Kaizik M.A., Hancock M., Moseley A.M., H.RD. DiTA: Diagnostic Test Accuracy database. <http://www.dita.org.au>. Accessed 6.03.24.
38. Kaizik M.A., Hancock M., Moseley A.M., Herbert R.D. DiTA: diagnostic Test Accuracy database. Indexing criteria and codes. <https://dita.org.au/learn/criteria-for-indexing/>. Accessed 6 March 2024.
39. Physiotherapy Evidence Database (PEDro). PEDro scale. <https://www.pedro.org.au/english/downloads/pedro-scale/>. Accessed 6.03.24.
40. Macedo LG, Elkins MR, Maher CG, Moseley AM, Herbert RD, Sherrington C. There was evidence of convergent and construct validity of physiotherapy evidence database quality scale for physiotherapy trials. *J Clin Epidemiol.* 2010;63(8):920–925.
41. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther.* 2003;83(8):713–721.
42. Yamato TP, Maher C, Koes B, Moseley A. The PEDro scale had acceptably high convergent validity, construct validity, and interrater reliability in evaluating methodological quality of pharmaceutical trials. *J Clin Epidemiol.* 2017;86:176–181.
43. Hollingworth W, Medina LS, Lenkinski RE, et al. Interrater reliability in assessing quality of diagnostic accuracy studies using the QUADAS tool. A preliminary assessment. *Acad Radiol.* 2006;13(7):803–810.
44. Whiting PF, Rutjes AW, Dinnes J, Reitsma JB, Bossuyt PM, Kleijnen J. A systematic review finds that diagnostic reviews fail to incorporate quality despite available tools. *J Clin Epidemiol.* 2005;58(1):1–12.
45. Kaizik MA, Garcia AN, Hancock MJ, Herbert RD. Measurement properties of quality assessment tools for studies of diagnostic accuracy. *Braz J Phys Ther.* 2020;24(2):177–184.
46. Charnley J. Orthopaedic signs in the diagnosis of disc protrusion. With special reference to the straight-leg-raising test. *Lancet.* 1951;1(6648):186–192.
47. van den Hoogen HM, Koes BW, van Eijk JT, Bouter LM. On the accuracy of history, physical examination, and erythrocyte sedimentation rate in diagnosing low back pain in general practice. A criteria-based review of the literature. *Spine (Phila Pa 1976).* 1995;20(3):318–327.