



OPEN ACCESS

EDITED BY

Oleksandr P. Romanchuk,
Lesya Ukrainka Volyn National University,
Ukraine

REVIEWED BY

Dawn Gulick,
self-employed, United States

*CORRESPONDENCE

Shasha Wang,
✉ wss_rehabil@163.com
Jia Han,
✉ jia.han@canberra.edu.au

RECEIVED 12 July 2025

ACCEPTED 22 August 2025

PUBLISHED 09 September 2025

CITATION

Luan L, Ji W, Ganderton C, Farragher J,
Pappas E, Adams RD, Wang S and Han J (2025) Is
the anterior drawer test still valuable for
diagnosing mechanical ankle instability in
clinical practice and research?
Front. Bioeng. Biotechnol. 13:1664779.
doi: 10.3389/fbioe.2025.1664779

COPYRIGHT

© 2025 Luan, Ji, Ganderton, Farragher, Pappas,
Adams, Wang and Han. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Is the anterior drawer test still valuable for diagnosing mechanical ankle instability in clinical practice and research?

Lijiang Luan^{1,2}, Wenxuan Ji³, Charlotte Ganderton^{3,4,5},
Joshua Farragher^{3,4}, Evangelos Pappas^{3,4}, Roger David Adams^{6,7},
Shasha Wang^{3*} and Jia Han^{1,3,6,8*}

¹Jiangwan Hospital of Hongkou District, Shanghai, China, ²School of Physical Education and Health, Sanming University, Sanming, China, ³College of Rehabilitation Sciences, Shanghai University of Medicine and Health Sciences, Shanghai, China, ⁴School of Health and Biomedical Sciences, RMIT University, Melbourne, VIC, Australia, ⁵Department of Nursing and Allied Health, Swinburne University of Technology, Melbourne, VIC, Australia, ⁶Research Institute for Sport and Exercise, University of Canberra, Canberra, ACT, Australia, ⁷School of Health Sciences, The University of Sydney, Sydney, NSW, Australia, ⁸Tongji University, Shanghai, China

The anterior drawer test (ADT) has been widely used in assessing mechanical ankle instability (MAI), yet its applicability has been questioned. This study aims to clarify the clinical value of the ADT. Five databases were searched in January 2025. Studies investigating the properties of the ADT were included. Data on reliability, validity, diagnostic accuracy, and responsiveness were extracted. A total of 424 studies were screened, and 45 studies were included. The ADT generally demonstrated good to excellent intra-rater reliability, but relatively poor inter-rater reliability. Criterion validity was supported by moderate to strong correlations with imaging and anatomical measurements. The diagnostic accuracy of ADT and its variants, such as the anterolateral and reverse ADTs, showed to be generally favorable. Inconsistencies in responsiveness across studies highlighted the need for population-specific classifications and the urgent establishment of corresponding measurement standards. The manual ADT exhibited a minimal detectable change value of 1.995 mm for intra-rater reliability. Overall, the ADT provides valuable insights for diagnosing MAI, but its inter-rater reliability and accuracy may be influenced by examiners' clinical experience and testing methods. Standardized protocols and advanced grading systems are needed to minimize inter-rater variability and enhance its result consistency, precision, and clinical utility.

KEYWORDS

anterior drawer test, mechanical ankle instability, reliability, validity, minimal detectable change

Introduction

The stability of the ankle joint plays a crucial role in maintaining balance and lower limb movement functions (Brown et al., 2015). While neuromuscular control is vital, ankle stability significantly depends on the structural integrity of the joint itself (Kirby et al., 2005). Ankle sprains can lead to varying degrees of ligamentous damage, potentially resulting in chronic ankle instability (Kerkhoffs et al., 2012; Lindstrand and Mortensson, 1977). Accurate diagnosis of the ligamentous integrity and mechanical

stability is crucial for formulating management strategies, developing training regimens, and assessing intervention effectiveness (Vuurberg et al., 2018).

Clinically, ankle stability is typically assessed using three primary methods. The first involves imaging modalities, often considered the gold standard, such as ultrasound, magnetic resonance imaging, and stress radiography (Kerkhoffs et al., 2012; Vuurberg et al., 2018). These techniques provide relatively accurate assessments of joint laxity or ligamentous structural integrity, but they are often time-consuming and costly (Kirby et al., 2005; Chang et al., 2021; Park, 2021). The second method is direct observation via surgery, such as arthroscopy; although this may be accurate, it is invasive and expensive (Vuurberg et al., 2018; Chang et al., 2021). The third and more commonly utilized approach is manual testing, which provides a timely and economical means of assessing ankle stability (Sun et al., 2018; Netterstrom-Wedin et al., 2022; Beynon et al., 2022).

Despite the convenience of manual testing, concerns about its reliability, validity, and diagnostic accuracy persist (Park, 2021; Netterstrom-Wedin et al., 2022; Beynon et al., 2022). These factors are critical for evaluating a diagnostic tool's effectiveness and align closely with the COSMIN guidelines, which provide core standards for assessing measurement tool quality (Mokkink et al., 2010). By investigating these aspects, one can gain a comprehensive understanding of the strengths and weaknesses of the ADT, offering clearer insights into its results and providing valuable guidance for health professionals when selecting diagnostic tools (Netterstrom-Wedin et al., 2022; Beynon et al., 2022; Mokkink et al., 2010). A diagnostic tool built on solid retest properties ensures reliable and valid measurements, supporting its use in clinical practice and research (Vuurberg et al., 2018; Chang et al., 2021).

Despite its prevalence, a systematic evaluation of the ADT remains lacking, and debates about its efficacy and accuracy persist (Netterstrom-Wedin et al., 2022). Therefore, this review aims to investigate the ADT, providing evidence-based guidance on its applicability as a diagnostic tool for assessing ankle instability.

Methods

This study has been registered in the International Prospective Register of Systematic Reviews (PROSPERO) database (registration number: CRD42024585466).

Search strategy

A systematic search was conducted in January 2025 using PubMed, Embase, Cochrane Library, Web of Science, and EBSCO databases to identify studies involving the use of ADT for the ankle joint. The search strategy focused on the terms 'ankle,' 'ADT,' and characteristics related to the reliability and retest evaluation of the drawer test methods (Supplementary Material S1A). Additionally, studies related to clinical examination of the ankle joint were screened, and relevant references were reviewed to identify additional studies for inclusion.

Study selection

The studies obtained from the systematic search and those included in relevant reviews were screened. The titles and abstracts of identified studies were initially screened to exclude ineligible studies. After the initial screening, eligibility was further assessed through full-text review based on the following inclusion criteria: (1) the ADT was performed on the ankle joint; (2) the evaluation results of the ADT were reported; (3) at least one psychometric property of the ADT was reported. Exclusion criteria: animal experiments, simulators or prosthetics, non-experimental detection surveys, case reports, reviews, and clinical trial registrations.

Quality assessment

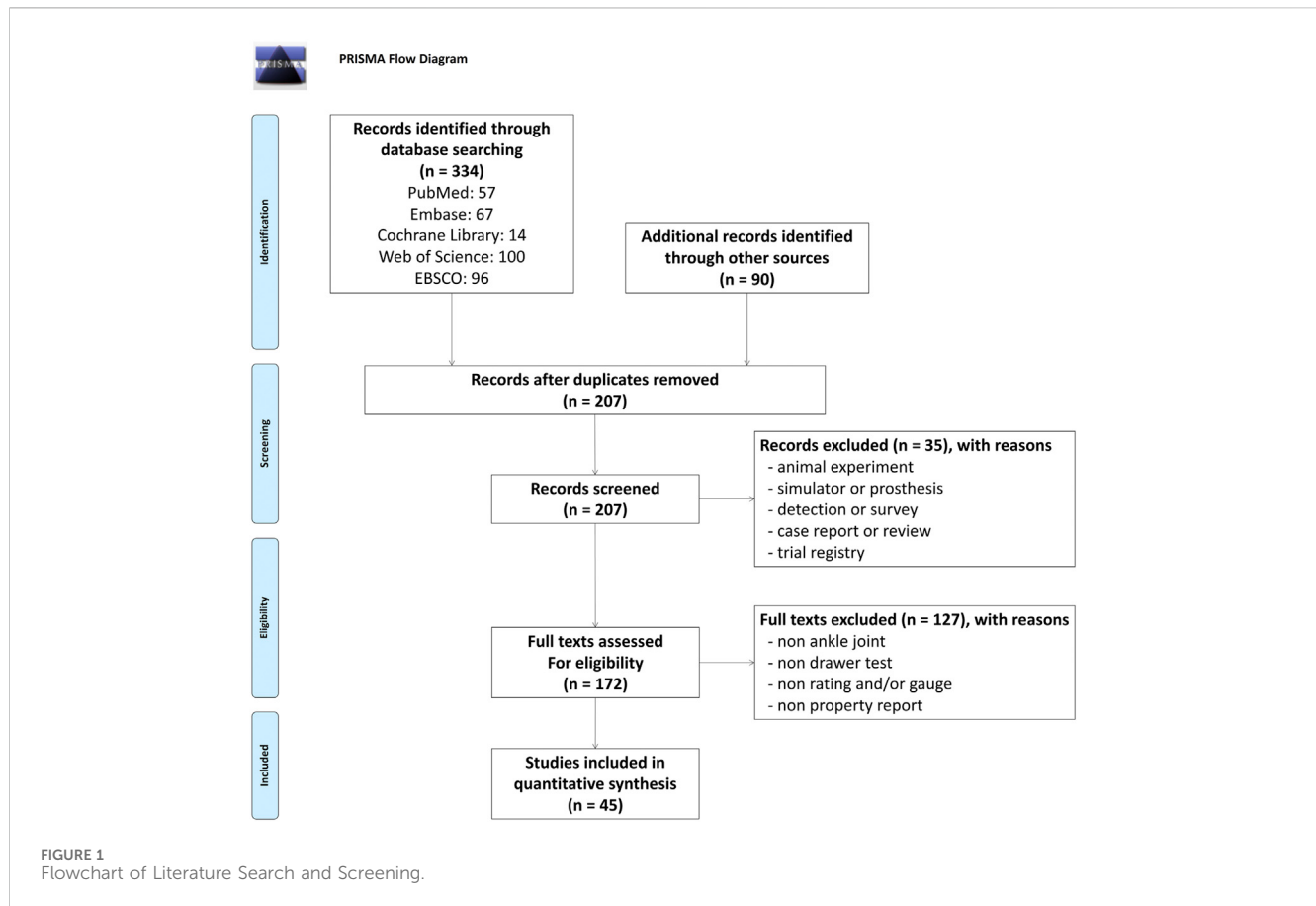
The quality of the studies that met the inclusion criteria was assessed using the Critical Appraisal Tool (CAT) (Brink and Louw, 2012), which is based on the QUADAS and QAREL (Whiting et al., 2003; Lucas et al., 2010). The CAT is scientifically robust, and its details are provided in Supplementary Material S1B.

Outcome measures

This study investigated the reliability and validity of the ADT, which are key components in the evaluation of clinical diagnostic tools (Beynon et al., 2022). Assessing the reliability and validity of the ADT is crucial for ensuring accurate test results and robust statistical conclusions (Beynon et al., 2022). Additionally, diagnostic accuracy and responsiveness were also key outcomes in this study, as the evaluation of these two aspects can measure the practical utility of the anterior drawer test, thereby reflecting its applicability (Sman et al., 2013). Specifically, regarding reliability, the ADT does not involve multiple items, so the outcome of internal consistency was excluded, with the focus mainly on test-retest reliability (intra/inter-rater). Regarding validity, this study did not delve into content validity, based on subjective evaluations; rather, it primarily analyzed construct validity and criterion validity. A focus was placed on specificity and sensitivity as they sufficiently reflect diagnostic accuracy; indicators such as likelihood ratio, predictive value, and accuracy, were not elaborated on, which can be derived through calculations (Netterstrom-Wedin et al., 2022; Sman et al., 2013). Regarding responsiveness, considering the need to provide some reference value for clinical evaluation, this study primarily reviewed the displacement results of the ADT.

Data extraction

Data on eligible studies were extracted, including demographic information and diagnostic conditions. To focus on the measurement effectiveness of the ADT, this study only considered participants who had undergone the ADT. Examiner experience performing the ADT, implementation measures, testing modes, and the evaluation methods used were summarized. Further,



specific data on reliability, validity, diagnostic accuracy, responsiveness are provided in [Supplementary Material S1C](#).

Exploration of the minimal detectable change

The Minimal Detectable Change (MDC) represents the smallest measurable difference that exceeds the measurement error of an assessment tool, indicating true clinical change rather than random error or variability ([Vuurberg et al., 2018; Netterstrom-Wedin et al., 2022](#)). MDC is calculated using the Standard Error of Measurement (SEM), quantifying the tool's inherent measurement error. MDC reflects the tool's sensitivity in detecting clinically meaningful changes, serving as a crucial reference in both clinical practice and research ([Sman et al., 2013; Phisitkul et al., 2009; Croy et al., 2013](#)). Therefore, exploring the MDC values in ADT is of great significance and key guiding role, and the specific methods adopted in this study are described in [Supplementary Material S1D](#).

Results

Literature search and screening

A total of 424 studies were identified through database searches and references from related reviews. After removing duplicate

studies, the titles or abstracts of 207 studies were screened. Subsequently, 172 full-text studies were reviewed, with 127 excluded for not addressing the ankle ADT or relevant properties. Ultimately, 45 studies (3,474 participants) met the inclusion criteria for this review ([Phisitkul et al., 2009; Croy et al., 2013; Gomes et al., 2018; Parasher et al., 2012; Vaseenon et al., 2012; George et al., 2020; Hosseinian et al., 2022; Großterlinden et al., 2016; Wilkin et al., 2012; van Dijk et al., 1996a; Li et al., 2020; Wiebking et al., 2015; Spahn, 2004; Wenning et al., 2021; van Dijk et al., 1996b; Chen et al., 2022; Lin et al., 2013; Raatikainen et al., 1992; van den Hoogenband et al., 1978; Prins, 1978; Lindstrand, 1976; Funder et al., 1988; Cho et al., 2016; Chandnani et al., 1994; Ahovuo et al., 1988; Rijke and Vierhout, 1990; Blanshard et al., 1986; Johannsen, 1978; Beumer et al., 2002; Yokoe et al., 2023; Yokoe et al., 2022; Azni et al., 2020; Chen et al., 2023; Iwata et al., 2023; Sillevs et al., 2022; Song et al., 2021; Murahashi et al., 2023; Teramoto et al., 2021; Kataoka et al., 2022; Yokoe et al., 2021; Kawabata et al., 2023; Docherty and Rybak-Webb, 2009; Saengsin et al., 2022; Gulick, 2024; Iwata et al., 2024](#)). The selection process and reasons for exclusion are illustrated in [Figure 1](#).

Quality assessment of included studies

Most included studies that assessed the reliability and/or validity of the ankle ADT demonstrated acceptable methodological rigor;

however, several limitations were identified. The detailed results of the quality assessment are summarized in [Supplementary Material S1E](#).

Characteristics of included studies

Among the eligible studies, most included participants with an ankle injury. The qualifications of the examiners performing the ADT varied. Most studies focused on assessing the anterior talofibular ligament (ATFL), with some studies also including the calcaneofibular ligament (CFL). The drawer test methods included traditional anterior drawer test (TADT), anterolateral drawer test (ALDT), reverse anterior drawer test (RADT), reverse anterolateral drawer test (RALDT), and instrumented anterior drawer test (IADT). The most common evaluation method was the subjective judgment by the examiner, with some studies using imaging as a reference for rating. Diagnostic accuracy studies primarily used imaging as the reference standard. The specific content is detailed in [Supplementary Material S1F](#).

Properties of ADTs

Despite inconsistencies in study populations, examiners, and experimental conditions, the reliability of various ADT modes was generally satisfactory for intra-rater assessment, as reflected by ICC and Kappa values ([Parasher et al., 2012](#); [Kataoka et al., 2022](#)). In contrast, most inter-rater reliability results were poor to moderate, only one or two studies reached high values ([Chen et al., 2022](#); [Saengsin et al., 2022](#)). Factors like examiner experience can affect inter-rater reliability, highlighting the need for careful assessment protocols. Three studies reported lower reliability ([Wilkin et al., 2012](#); [Li et al., 2020](#); [Beumer et al., 2002](#)), primarily due to small sample sizes and atypical testing positions (e.g., supine). Overall, the ADT and its variants (ALDT, RADT, IADT) generally maintain moderate to excellent reliability.

In terms of criterion validity, ADT generally showed moderate to strong positive correlation with imaging modalities such as ultrasound, magnetic resonance imaging, and stress radiography ([Gomes et al., 2018](#); [Saengsin et al., 2022](#)). However, research on construct validity remains limited, with three studies showing low correlation between ADT and certain scales such as CAIT, VAS, and Beighton ([Wilkin et al., 2012](#); [Spahn, 2004](#); [Yokoe et al., 2022](#)), while two studies indicating high correlation between ADT and CAIT as well as Beighton score ([Wenning et al., 2021](#); [Song et al., 2021](#)). Further exploration of the subjective assessment correlation is necessary due to this contradiction.

ADT's sensitivity and specificity varied across studies depending on parameters associated with the condition of ankle instability and examiner experience. Good specificity was observed in patients with acute supination trauma, while sensitivity was often poor in those reporting chronic ankle instability ([Gomes et al., 2018](#); [Funder et al., 1982](#)). Parameters such as duration of instability and ligament injury degree impact diagnostic accuracy. One study found that senior and junior raters showed 80% and 40% sensitivity for the ADT, respectively ([Murahashi et al., 2023](#)).

The responsiveness of ADT varied across studies, with significant differences in displacement values between patient groups. For instance, one study reported a mean displacement of 4.5 mm for senior examiners vs. 3.26 mm for junior examiners ([Murahashi et al., 2023](#)). Specifically, mean displacement of 3.2 mm in injured ATFL and 0.6 mm in intact ATFL were observed ([Iwata et al., 2024](#)). Significant displacement differences were noted in patients with CAI or ligament injury, yet within-group variability remained high ([Saengsin et al., 2022](#)).

Detailed data above are presented in [Supplementary Material S2A–D](#), respectively.

The minimal detectable change of the ADT

Three studies reported both ICC and SEM values for reliability outcomes ([Parasher et al., 2012](#); [Wilkin et al., 2012](#); [Docherty and Rybak-Webb, 2009](#)). Additionally, one study provided ICC (intra) and standard deviation (SD) for each rater's measurements assessing intra-rater reliability ([Kataoka et al., 2022](#)). Calculations indicated that manual ADT testing exhibited a maximal MDC of 1.995 mm for intra-rater reliability, while the MDC for instrumented ADT testing was higher, at 6.153 mm. Furthermore, discrepancies between raters in ADT application showed a maximal MDC of 6.291 mm or 4.684 units on an 8-point Likert scale. Detailed results are presented in [Supplementary Material S2E](#).

Discussion

While certain limitations exist, such as inter-rater reliability issues and subjective grading, the ADT remains a practical, cost-effective, and widely accepted tool in clinical practice. By synthesizing insights from reliability and validity data, sensitivity and specificity reports, and MDC analyses, this discussion underscores the ADT's continued relevance and identifies areas for refinement, including standardized protocols and enhanced grading systems.

The properties of the ADT

As noted, the reliability of the ADT is generally high, with most studies reporting moderate to excellent intra-rater and inter-rater reliability. However, several studies highlight that test reliability is significantly affected by factors such as examiner experience, test variations, and subject positioning. Specifically, intra-rater reliability is generally high, particularly when the same examiner conducts the test multiple times. However, differences are evident when examining the same test performed by different raters, indicating that examiner experience can impact on reliability. Additionally, ADT's reliability was lower when using multiple variations of the test (e.g., ALDT, RADT) compared to the traditional approach ([Li et al., 2020](#); [Beumer et al., 2002](#)), suggesting that more complex testing variations may increase error and variability. Therefore, proper training and uniform test protocols are essential to minimize discrepancies and enhance the consistency of results.

The validity of the ADT is primarily assessed by comparing its results with imaging and surgical findings. Most studies demonstrate moderate to strong criterion validity through correlations with imaging modalities, confirming that ADT results align reasonably well with imaging assessments. However, construct validity remains less well-explored, and the limited evidence mainly concerns rating scales. For example, the Beighton scale, commonly used to evaluate generalized joint hypermobility by assessing the range of motion in the little finger, thumb, elbow, knee and trunk, which is a 9-point clinical scoring system (Yokoe et al., 2022; Song et al., 2021). In this review, a correlation of $r = 0.719$ between Beighton scores and ADT was found (Song et al., 2021), indicating a strong relationship in some populations, while Beighton scores did not demonstrate significant correlations with ADT results in the other two studies (Yokoe et al., 2023; Yokoe et al., 2022). This discrepancy highlights that ankle instability as measured by ADT may not be fully captured by generalized joint laxity scores, reinforcing the need for a more comprehensive assessment method that incorporates both global and local assessment.

The diagnostic accuracy of ADT is generally favorable but varies based on factors like examiner experience, test timing, and patient characteristics. For instance, two studies revealed stark differences between novice and experienced raters (Li et al., 2020; Murahashi et al., 2023), highlighting how examiner proficiency impacts test outcomes. As variations of the drawer test, such as ALDT, RADT, and IADT, experience less impact from examiner experience, potentially making them better suited for novices. Test timing also influences diagnostic outcomes; certain studies found that sensitivity and specificity differed significantly between 48 h and 5 days after injury (van Dijk et al., 1996a; van Dijk et al., 1996b). These findings underscore the importance of timing in diagnostic accuracy, suggesting ADT is more reliable after the acute inflammatory phase has subsided. Patient characteristics, like injury severity and the acute vs. chronic nature of the instability, can also affect diagnostic accuracy. For instance, two studies demonstrated lower accuracy in patients with chronic syndesmotic ruptures (Johannsen, 1978; Beumer et al., 2002), likely due to injury complexity. Additionally, higher diagnostic accuracy was found in patients with severe injuries (Raatikainen et al., 1992), suggesting that ADT may be more effective in detecting more pronounced mechanical instability. Therefore, ADT's diagnostic accuracy is highly dependent on contextual factors such as examiner expertise, timing of conduction in relation to injury and injury severity.

The responsiveness of ADT varies across studies, likely influenced by examiner experience and patient demographics, including injury severity and time since injury. Differences in displacement between senior and junior examiners were noted (Murahashi et al., 2023), indicating that examiner proficiency impacts test responsiveness. A study observed a difference between injured and intact ATFL (Iwata et al., 2024), suggesting acute injuries produce more pronounced displacement, whereas chronic cases may show less sensitivity to ADT. This indicates that the measurement of ADT is likely more effective for acute injury than for chronic condition. Displacement differences were also noted across symptomatic conditions; chronic injuries and multi-ligament tears (ATFL + CFL) showed larger displacements compared to isolated ATFL

injuries (Saengsin et al., 2022). These differences indicate that ADT's responsiveness is influenced by both examiner experience and injury severity, which affects ligamentous stability and displacement readings. Thus, establishing population-specific MDC and MCID values is essential for improving ADT's responsiveness in clinical settings.

The minimal detectable change for the ADT and additional considerations

In this study, the MDC for the ADT was explored as an essential measure of the smallest clinically relevant change that exceeds measurement error. Based on the MDC values and findings across studies (Supplementary Material S2E), we can derive several key considerations for improving the precision and clinical applicability of the ADT.

When performing ADT, even experienced raters may show considerable error between their results. This indicates variability exists between raters, even among experienced examiners. Therefore, when multiple raters are involved in ADT assessments, it is crucial to be cautious with the analysis of the results, as discrepancies can arise.

Manual ADT testing may demonstrate smaller errors compared to instrumented ADT. This could be due to the inherent stability of manual testing, where the rater's movements are typically controlled until no further motion is observed (Parasher et al., 2012; Wilkin et al., 2012; Docherty and Rybak-Webb, 2009). In contrast, instrumented ADT often operates with a fixed force, such as 80 N, 120 N, or 150 N (Chen et al., 2022; Kataoka et al., 2022). However, due to the elastic nature of joint tissues, the sliding distance in instrumented ADT may vary as soft tissues adapt, leading to larger measurement errors (Wiebking et al., 2015; Chen et al., 2022; Murahashi et al., 2023). On the other hand, in manual ADT, the same maximum sliding distance is more consistently achieved with each assessment, resulting in smaller error margins (Parasher et al., 2012; Wilkin et al., 2012).

Using imaging devices, such as fluoroscopic images, results in smaller measurement errors and greater precision compared to tools like goniometers or arthrometers (Parasher et al., 2012; Saengsin et al., 2022). Although imaging provides a clearer and more accurate method of measuring, it is a more costly and time-consuming approach (Parasher et al., 2012; Sillevius et al., 2022). This trade-off must be considered when determining the most appropriate measurement method for clinical or research purposes.

The common grading system for ADT evaluations typically uses three levels (1, 2, 3), which are rather broad and lack the precision needed to distinguish subtle differences (Parasher et al., 2012; Vaseenon et al., 2012). For instance, the difference between level 2 and level 3 may not be significant enough to exceed the MDC, leading to potential misclassification. Additionally, using just three to four grading levels makes it difficult to effectively and accurately assess measurement errors. It is suggested that future ADT evaluations use a finer scale, such as an 8-point Likert scale, to improve precision in detecting measurement errors and enhance the validity of MDC estimation.

Recommendations for improving the ADT

Based on the characteristics, current applications, and identified limitations of the ADT discussed above, several recommendations for improvement are proposed to enhance its clinical utility.

While instrumented and imaging-based ADT methods offer measurable advancements in precision, they are not without limitations. Instrumented drawer tests apply fixed forces, which fail to account for individual differences such as foot length, body weight, or joint elasticity, which makes it challenging to achieve reliable results in diverse populations. Similarly, imaging-based methods, while providing precise displacement values, are often impractical in clinical settings due to their cost, inefficiency, and unclear clinical significance of the measurements.

In contrast, manual drawer testing remains a highly adaptable and patient-centered approach. Skilled therapists can adjust force application based on the patient's individual characteristics, minimizing variability caused by population differences. Furthermore, manual tests are quick, cost-effective, and well-suited to clinical practice. These qualities underscore the importance of continuing to rely on manual drawer tests while addressing their inherent limitations.

Despite its practicality, manual drawer testing has shortcomings. Variability among raters, as highlighted by differences in intra-rater and inter-rater reliability, presents a significant challenge. MDC intra-rater values tend to be lower and more consistent compared to inter-rater MDC values, reflecting the added variability when multiple raters are involved. Some studies demonstrated noticeable disparities in reliability across different raters (Parasher et al., 2012; Wilkin et al., 2012).

To enhance the consistency of manual testing, several recommendations are proposed: (a) Standardized protocols: Establish comprehensive and highly detailed operating procedures for ADT, including standardized verbal instructions and consistent force application techniques; also, current studies provide limited reporting on the position of the ankle in testing (dorsiflexion, neutral, plantarflexion), making deeply comparisons difficult. However, different testing positions may exert varying influences on the ATF and CF ligaments, highlighting the need for future research to standardize the specific ankle position during ADT implementation. (b) Training and certification: Implement rigorous training programs to ensure that examiners achieve a high level of proficiency before clinical application. Practical assessments and certification processes should be developed to evaluate the competency of testers, aligning their skills to a consistent standard. (c) Continuous monitoring: Encourage single-rater assessments for individual patients whenever feasible, as this minimizes inter-rater variability and strengthens reliability over the course of treatment. (d) Transition to a more granular grading scale, such as a 0–10 Likert scale or a modified VAS scale. These systems can offer enhanced precision, reflecting a continuum from “no instability” to “complete instability”. Expanding the ADT scoring system from the traditional 0–3/4 grades to a broader 0–8 or 0–10 scale may offer notable advantages. First, it increases the sensitivity of the assessment by allowing finer gradations of anterior talar translation, thereby capturing subtle variations in ankle laxity that might be overlooked with a narrower scale. Second, a broader scoring

range enhances clinical and research applicability by improving the ability to discriminate between patients with different levels of instability, reducing ceiling effects, and enabling more precise comparisons of therapeutic outcomes. (e) Develop population-specific MDC and MCID thresholds to better guide clinical decision-making for subgroups like patients with structural ankle instability. This refinement would enhance the accuracy and clinical utility of manual ADT, facilitating more precise diagnoses and targeted interventions.

These recommendations aim to bridge the gap between manual and advanced ADT methods, ensuring that manual drawer testing remains a cornerstone of ankle instability assessment while evolving to meet the demands of modern clinical and research practices. Study limitations are presented in [Supplementary Material S2F](#).

Author contributions

LL: Writing – original draft, Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization. WJ: Conceptualization, Formal Analysis, Methodology, Validation, Writing – review and editing. CG: Conceptualization, Formal Analysis, Methodology, Validation, Writing – review and editing. JF: Conceptualization, Formal Analysis, Methodology, Validation, Writing – review and editing. EP: Conceptualization, Formal Analysis, Methodology, Validation, Writing – review and editing. RA: Conceptualization, Formal Analysis, Methodology, Validation, Writing – review and editing. SW: Conceptualization, Formal Analysis, Methodology, Validation, Writing – review and editing. JH: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure

accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that

may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2025.1664779/full#supplementary-material>

References

- Ahovuo, J., Kaartinen, E., and Slati, P. (1988). Diagnostic value of stress radiography in lesions of the lateral ligaments of the ankle. *Acta radiol.* 29 (6), 711–714. doi:10.1177/028418518802900620
- Azmi, K. K., Saini, S. M., Nasir, N., and Bahari, N. (2020). Correlation of clinical examination, magnetic resonance imaging and surgical findings in diagnosing ankle joint ligament injuries. *Eur. J. Mol. Clin. Med.* 7 (2), 1474–1484.
- Beumer, A., Swierstra, B. A., and Mulder, P. G. (2002). Clinical diagnosis of syndesmotank ankle instability: evaluation of stress tests behind the curtains. *Acta Orthop. Scand.* 73 (6), 667–669. doi:10.1080/000164702321039642
- Beynon, A., Le May, S., and Theroux, J. (2022). Reliability and validity of physical examination tests for the assessment of ankle instability. *Chiropr. Man. Ther.* 30 (1), 58. doi:10.1186/s12998-022-00470-0
- Blanshard, K. S., Finlay, D. B., Scott, D. J., Ley, C. C., Siggins, D., and Allen, M. J. (1986). A radiological analysis of lateral ligament injuries of the ankle. *Clin. Radiol.* 37 (3), 247–251. doi:10.1016/s0009-9260(86)80328-2
- Brink, Y., and Louw, Q. A. (2012). Clinical instruments: reliability and validity critical appraisal. *J. Eval. Clin. Pract.* 18 (6), 1126–1132. doi:10.1111/j.1365-2753.2011.01707.x
- Brown, C. N., Rosen, A. B., and Ko, J. (2015). Ankle ligament laxity and stiffness in chronic ankle instability. *Foot and ankle Int.* 36 (5), 565–572. doi:10.1177/1071100714561057
- Chandnani, V. P., Harper, M. T., Ficke, J. R., Gagliardi, J. A., Rolling, L., Christensen, K. P., et al. (1994). Chronic ankle instability: evaluation with MR arthrography, MR imaging, and stress radiography. *Radiology* 192 (1), 189–194. doi:10.1148/radiology.192.1.8208935
- Chang, S. H., Morris, B. L., Saengsin, J., Tourné, Y., Guillo, S., Guss, D., et al. (2021). Diagnosis and treatment of chronic lateral ankle instability: review of our biomechanical evidence. *J. Am. Acad. Orthop. Surg.* 29 (1), 3–16. doi:10.5435/jaas-d-20-00145
- Chen, Y., Cao, S., Wang, C., Ma, X., and Wang, X. (2022). Quantitative analysis with load-displacement ratio measured via digital arthrometer in the diagnostic evaluation of chronic ankle instability: a cross-sectional study. *J. Orthop. Surg. Res.* 17 (1), 287. doi:10.1186/s13018-022-03177-3
- Chen, Y., Cao, S., Wang, C., Zhang, C., Huang, J., Wang, X., et al. (2023). Diagnostic evaluation of mechanical ankle instability by comparing injured and uninjured contralateral ankles using arthrometry. *J. Athl. Train.* 58 (11-12), 934–940. doi:10.4085/1062-6050-0494.22
- Cho, J. H., Lee, D. H., Song, H. K., Bang, J. Y., Lee, K. T., and Park, Y. U. (2016). Value of stress ultrasound for the diagnosis of chronic ankle instability compared to manual anterior drawer test, stress radiography, magnetic resonance imaging, and arthroscopy. *Knee Surg. sports traumatology, Arthrosc. official J. ESSKA* 24 (4), 1022–1028. doi:10.1007/s00167-015-3828-9
- Croy, T., Kopenhagen, S., Saliba, S., and Hertel, J. (2013). Anterior talocrural joint laxity: diagnostic accuracy of the anterior drawer test of the ankle. *J. Orthop. Sports Phys. Ther.* 43 (12), 911–919. doi:10.2519/jospt.2013.4679
- Docherty, C. L., and Rybak-Webb, K. (2009). Reliability of the anterior drawer and talar tilt tests using the LigMaster joint arthrometer. *J. Sport Rehabil.* 18 (3), 389–397. doi:10.1123/jsr.18.3.389
- Funder, V., Jorgensen, J. P., Andersen, A., Andersen, S. B., Lindholmer, E., Niedermann, B., et al. (1982). Ruptures of the lateral ligaments of the ankle. Clinical diagnosis. *Acta Orthop. Scand.* 53 (6), 997–1000. doi:10.3109/17453678208992860
- George, J., Jaafar, Z., Hairi, I. R., and Hussein, K. H. (2020). The correlation between clinical and ultrasound evaluation of anterior talofibular ligament and calcaneofibular ligament tears in athletes. *J. Sports Med. Phys. Fit.* 60 (5), 749–757. doi:10.23736/s0022-4707.20.10050-1
- Gomes, J. L. E., Soares, A. F., Bastiani, C. E., and de Castro, J. V. (2018). Anterolateral talar palpation: a complementary test for ankle instability. *Foot ankle Surg. official J. Eur. Soc. Foot Ankle Surg.* 24 (6), 486–489. doi:10.1016/j.fas.2017.05.006
- Großterlinden, L. G., Hartel, M., Yamamura, J., Schoennagel, B., Bürger, N., Krause, M., et al. (2016). Isolated syndesmotank injuries in acute ankle sprains: diagnostic significance of clinical examination and MRI. *Knee Surg. sports traumatology, Arthrosc. official J. ESSKA* 24 (4), 1180–1186. doi:10.1007/s00167-015-3604-x
- Gulick, D. T. (2024). Using an arthrometer to quantify ankle laxity: an infographic summary. *Sports health* 16 (1), 26–28. doi:10.1177/19417381221147511
- Hosseinian, S. H. S., Aminzadeh, B., Rezaeian, A., Jarahi, L., Naeini, A. K., and Jangui, P. (2022). Diagnostic value of ultrasound in ankle sprain. *J. foot ankle Surg.* 61, 305–309. doi:10.1053/j.jfas.2021.08.008
- Iwata, S., Yamaguchi, S., Kimura, S., Hattori, S., Sasahara, J., Akagi, R., et al. (2023). Do standardization of the procedure reduce measurement variability of the sonographic anterior drawer test of the ankle? *J. Orthop. Sci.* 29, 1438–1444. doi:10.1016/j.jos.2023.11.006
- Iwata, S., Yamaguchi, S., Kimura, S., Hattori, S., Mikami, Y., Kawasaki, Y., et al. (2024). Variability in sonographic anterior drawer test measurements of the ankle: experienced versus beginner examiners. *J. Orthop. Sci.* 29 (1), 243–248. doi:10.1016/j.jos.2022.12.005
- Johannsen, A. (1978). Radiological diagnosis of lateral ligament lesion of the ankle. A comparison between talar tilt and anterior drawer sign. *Acta Orthop. Scand.* 49 (3), 295–301. doi:10.3109/17453677809005768
- Kataoka, K., Hoshino, Y., Nagamune, K., Nukuto, K., Yamamoto, T., Yamashita, T., et al. (2022). The quantitative evaluation of anterior drawer test using an electromagnetic measurement system. *Sports Biomech.* 21 (4), 550–561. doi:10.1080/14763141.2021.1918754
- Kawabata, M., Kumazawa, Y., Takagi, K., Okada, H., Miyatake, K., Kobayashi, T., et al. (2023). Reliability and validity of ultrasonographic automated length measurement system for assessing talofibular anterior instability in acute lateral ankle sprain. *Sci. Rep.* 13 (1), 3098. doi:10.1038/s41598-023-30079-z
- Kerkhoffs, G. M., van den Bekerom, M., Elders, L. A., van Beek, P. A., Hulleger, W. A. M., Bloemers, G. M. F. M., et al. (2012). Diagnosis, treatment and prevention of ankle sprains: an evidence-based clinical guideline. *Br. J. Sports Med.* 46 (12), 854–860. doi:10.1136/bjsports-2011-090490
- Kirby, A. B., Beall, D. P., Murphy, M. P., Ly, J. Q., and Fish, J. R. (2005). Magnetic resonance imaging findings of chronic lateral ankle instability. *Curr. Probl. Diagn. Radiol.* 34 (5), 196–203. doi:10.1067/j.cpradiol.2005.06.003
- Li, Q., Tu, Y., Chen, J., Shan, J., Yung, P. S. H., Ling, S. K. K., et al. (2020). Reverse anterolateral drawer test is more sensitive and accurate for diagnosing chronic anterior talofibular ligament injury. *Knee Surg. sports traumatology, Arthrosc. official J. ESSKA* 28 (1), 55–62. doi:10.1007/s00167-019-05705-x
- Lin, C. Y., Shau, Y. W., Wang, C. L., Chai, H. M., and Kang, J. H. (2013). Quantitative evaluation of the viscoelastic properties of the ankle joint complex in patients suffering from ankle sprain by the anterior drawer test. *Knee Surg. sports traumatology, Arthrosc. official J. ESSKA* 21 (6), 1396–1403. doi:10.1007/s00167-013-2459-2
- Lindstrand, A. (1976). *Lateral lesions in sprained ankles: a clinical and roentgenological study with special reference to anterior instability of the talus*. Lund, Sweden: Lund University.
- Lindstrand, A., and Mortenson, W. (1977). Anterior instability in the ankle joint following acute lateral sprain. *Acta Radiol. Diagn Stockh.* 18 (5), 529–539. doi:10.1177/028418517701800504
- Lucas, N. P., Macaskill, P., Irwig, L., and Bogduk, N. (2010). The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *J. Clin. Epidemiol.* 63 (8), 854–861. doi:10.1016/j.jclinepi.2009.10.002
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., et al. (2010). 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.* 19, 539–549. doi:10.1007/s11136-010-9606-8
- Murahashi, Y., Teramoto, A., Takahashi, K., Okada, Y., Okimura, S., Imamura, R., et al. (2023). High reproducibility of a novel supported anterior drawer test for

- diagnosing ankle instability. *BMC Musculoskelet. Disord.* 24 (1), 148. doi:10.1186/s12891-023-06246-8
- Netterstrom-Wedin, F., Matthews, M., and Bleakley, C. (2022). Diagnostic accuracy of clinical tests assessing ligamentous injury of the talocrural and subtalar joints: a systematic review with meta-analysis. *Sports health* 14 (3), 336–347. doi:10.1177/19417381211029953
- Parasher, R. K., Nagy, D. R., Em, A. L., Phillips, H. J., and Mc Donough, A. L. (2012). Clinical measurement of mechanical ankle instability. *Man. Ther.* 17 (5), 470–473. doi:10.1016/j.math.2012.02.015
- Park, H.-w. (2021). Are magnetic resonance imaging findings of ankle instability always correlated with operative and physical examination findings? *J. Korean Foot Ankle Soc.* 25 (1), 1–5. doi:10.14193/jkf.2021.25.1.1
- Phisitkul, P., Chaichankul, C., Sripongsai, R., Prasitdamrong, I., Tengtrakulcharoen, P., and Suarachawatana, S. (2009). Accuracy of anterolateral drawer test in lateral ankle instability: a cadaveric study. *Foot and ankle Int.* 30 (7), 690–695. doi:10.3113/fai.2009.0690
- Prins, J. G. (1978). Diagnosis and treatment of injury to the lateral ligament of the ankle. A comparative clinical study. *Acta Chir. Scand. Suppl.* 486, 3–149.
- Raatikainen, T., Putkonen, M., and Puranen, J. (1992). Arthrography, clinical examination, and stress radiograph in the diagnosis of acute injury to the lateral ligaments of the ankle. *Am. J. sports Med.* 20 (1), 2–6. doi:10.1177/036354659202000102
- Rijke, A. M., and Vierhout, P. A. (1990). Graded stress radiography in acute injury to the lateral ligaments of the ankle. *Acta Radiol* 31 (2), 151–155. doi:10.3109/02841859009177479
- Saengsin, J., Bhimani, R., Sato, G., C. Hagemeyer, N., Mirochnik, K., Lubberts, B., et al. (2022). Use of portable ultrasonography for the diagnosis of lateral ankle instability. *J. Orthop. Res. official Publ. Orthop. Res. Soc.* 40 (10), 2421–2429. doi:10.1002/jor.25256
- Sillevis, R., Shamus, E., and van Duijn, A. (2022). Evaluation of anterotalofibular and calcaneofibular ligament stress tests utilizing musculoskeletal ultrasound imaging. *Physiother Theory Pract* 38(10), 1488–1498.
- Sman, A. D., Hiller, C. E., and Refshauge, K. M. (2013). Diagnostic accuracy of clinical tests for diagnosis of ankle syndesmosis injury: a systematic review. *Br. J. Sports Med.* 47 (10), 620–628. doi:10.1136/bjsports-2012-091702
- Song, J. H., Kang, C., Kim, N. S., Yi, J. W., Lee, G. S., Jang, M. G., et al. (2021). Evaluation of the uninjured anterior talofibular ligament by ultrasound for assessing generalized joint hypermobility. *Foot ankle Surg. official J. Eur. Soc. Foot Ankle Surg.* 27 (3), 256–262. doi:10.1016/j.fas.2020.10.014
- Spahn, G. (2004). The ankle meter: an instrument for evaluation of anterior talar drawer in ankle sprain. *Knee Surg. sports traumatology, Arthrosc. official J. ESSKA* 12 (4), 338–342. doi:10.1007/s00167-003-0477-1
- Sun, Y., Wang, H., Tang, Y., Qin, S., Zhao, M., and Zhang, F. (2018). Diagnosis and treatment of chronic lateral ankle instability with ligamentum bifurcatum injury: an observational study. *Med. Baltim.* 97 (9), e0028. doi:10.1097/md.00000000000010028
- Teramoto, A., Iba, K., Murahashi, Y., Shoji, H., Hirota, K., Kawai, M., et al. (2021). Quantitative evaluation of ankle instability using a capacitance-type strain sensor. *Foot and ankle Int.* 42 (8), 1074–1080. doi:10.1177/1071100721996714
- van den Hoogenband, C. R., van Moppes, F. I., Stapert, J. W., and Greep, J. M. (1978/1984). Clinical diagnosis, arthrography, stress examination and surgical findings after inversion trauma of the ankle. *Arch. Orthop. Trauma Surg.* 103 (2), 115–119. doi:10.1007/bf00389583
- van Dijk, C. N., Lim, L. S., Bossuyt, P. M., and Marti, R. K. (1996a). Physical examination is sufficient for the diagnosis of sprained ankles. *J. Bone Jt. Surg. Br.* 78 (6), 958–962. doi:10.1302/0301-620x.78b6.0780958
- van Dijk, C. N., Mol, B. W., Lim, L. S., Marti, R. K., and Bossuyt, P. M. (1996b). Diagnosis of ligament rupture of the ankle joint. Physical examination, arthrography, stress radiography and sonography compared in 160 patients after inversion trauma. *Acta Orthop. Scand.* 67 (6), 566–570. doi:10.3109/17453679608997757
- Vaseenon, T., Gao, Y., and Phisitkul, P. (2012). Comparison of two manual tests for ankle laxity due to rupture of the lateral ankle ligaments. *Iowa Orthop. J.* 32, 9–16.
- Vuurberg, G., Hoortje, A., Wink, L. M., van der Doelen, B. F. W., van den Bekerom, M. P., Dekker, R., et al. (2018). Diagnosis, treatment and prevention of ankle sprains: update of an evidence-based clinical guideline. *Br. J. Sports Med.* 52 (15), 956. doi:10.1136/bjsports-2017-098106
- Wenning, M., Gehring, D., Lange, T., Fuerst-Meroth, D., Streicher, P., Schmal, H., et al. (2021). Clinical evaluation of manual stress testing, stress ultrasound and 3D stress MRI in chronic mechanical ankle instability. *BMC Musculoskelet. Disord.* 22 (1), 198. doi:10.1186/s12891-021-03998-z
- Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., and Kleijnen, J. (2003). The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med. Res. Methodol.* 3, 25. doi:10.1186/1471-2288-3-25
- Wiebking, U., Pacha, T. O., and Jagodzinski, M. (2015). An accuracy evaluation of clinical, arthrometric, and stress-sonographic acute ankle instability examinations. *Foot ankle Surg. official J. Eur. Soc. Foot Ankle Surg.* 21 (1), 42–48. doi:10.1016/j.fas.2014.09.006
- Wilkin, E. J., Hunt, A., Nightingale, E. J., Munn, J., Kilbreath, S. L., and Refshauge, K. M. (2012). Manual testing for ankle instability. *Man. Ther.* 17 (6), 593–596. doi:10.1016/j.math.2012.03.007
- Yokoe, T., Tajima, T., Kawagoe, S., Yamaguchi, N., Morita, Y., and Chosa, E. (2021). The ratio of stress to nonstress anterior talofibular ligament length on ultrasonography: normative values. *Orthop. J. sports Med.* 9 (11), 23259671211056305. doi:10.1177/23259671211056305
- Yokoe, T., Tajima, T., Yamaguchi, N., Morita, Y., and Chosa, E. (2022). Association between the Beighton score and stress ultrasonographic findings of the anterior talofibular ligament in healthy young women: a cross-sectional study. *J. Clin. Med.* 11 (7), 1759. doi:10.3390/jcm11071759
- Yokoe, T., Tajima, T., Yamaguchi, N., Morita, Y., and Chosa, E. (2023). The anterior talofibular ligament ratio was greater in young men with generalized joint laxity than in those without generalized joint laxity. *Knee Surg. Sports Traumatol. Arthrosc. Official J. Esska.* 31 (5), 1994–2000. doi:10.1007/s00167-022-07109-w