INNOVAPATH™
The Premier Journal Bridging Discovery & Innovation

Review Article
https://doi.org/10.63501/hg85j716

Fairness Begins at the Gate: Why Pakistan's MDCAT Must Move Beyond Raw Scores

Sohail Rao, MD, MA, DPhil[1]

[1]The Texas Institute for Graduate Medical Education, 1917 Ashland Street, Houston, TX 77008.

[1]Corresponding Author Email: srao@tigme.org

ABSTRACT

Background: The Medical and Dental College Admission Test (MDCAT) is the primary gatekeeper for entry into medical and dental education in Pakistan. Recent controversies surrounding MDCAT 2023–2024 paper leaks, skewed score distributions, and the use of raw scores across multiple sessions and years have raised serious concerns about fairness, validity, and public trust. Court judgments and parliamentary debates have further highlighted weaknesses in governance and psychometric practice.

Objective: To critically evaluate the current use of raw scores in MDCAT across multiple test forms and years, compare this approach with international best practices in high-stakes testing, and propose a practical roadmap for psychometrically defensible and ethically robust reform.

Methods: This narrative review synthesizes evidence from Pakistani legal documents, public notices, and media reports on MDCAT 2022–2024, alongside international standards in educational measurement and technical documentation from major testing agencies (e.g., ETS, NCCAOM, NCBE, AERAs Standards for Educational and Psychological Testing).

Results: Findings indicate that: (1) MDCAT scores are currently reported and interpreted as raw percentages, while being declared valid for up to three years across cohorts under Section 17 of the PM&DC Act 2022; (2) multiple court proceedings and independent reviews have noted statistical anomalies, including very weak correlations between MDCAT and board marks, unusually high clusters of near-perfect scores, and credible allegations of paper leaks across sessions; and (3) international high-stakes examinations with multiple forms (e.g., MCAT, LSAT, GRE) rely on equating and scaled scores to ensure comparability, as required by professional standards.

Discussion: Using unadjusted raw scores from non-equated forms in a multi-session, multi-year, high-stakes exam is psychometrically indefensible and ethically problematic. It risks hidden advantage, systemic disadvantage, and erosion of trust. Pakistan can align MDCAT with global best practice by legally mandating equated and scaled scoring, creating an independent National Center for Health Professions Assessment, adopting anchor-item designs and item response theory (IRT), and enhancing transparency.

Conclusion: Scaling and equating are not statistical luxuries but instruments of justice. For a country where access to competent healthcare can mean life or death, fairness at the very first gate, who is allowed to become a doctor, is an ethical imperative.

Review Article
https://doi.org/10.63501/hg85j716

KEYWORDS: *MDCAT, Pakistan, admissions testing, score equating, scaling, fairness, psychometrics, PM&DC*

INTRODUCTION

Each year, tens of thousands of students in Pakistan sit for MDCAT with the expectation that this single, high-stakes examination will be scored fairly and consistently. MDCAT is mandated by the Pakistan Medical and Dental Council (PMDC) as the primary admissions test for MBBS and BDS programs, with scores now declared valid for up to three years under Section 17(4) of the PM&DC Act 2022.[1,2,12]

However, the 2023 and 2024 MDCAT administrations have been marred by allegations of paper leaks, inconsistent invigilation, unusual score distributions, clusters of near-perfect marks, and weak correlations between MDCAT and board examination results.[1,3,4,13–15] The Sindh High Court, for example, has noted that in one dataset the correlation between MDCAT scores and board marks was only about 0.16 and that groups of students from specific centers achieved improbably high scores, prompting stays of result announcements and independent inquiries.[1,3,4].

At the same time, the regulatory stance has emphasized that MDCAT raw scores, as declared, are final and may be used across multiple years without recalculation or "normalization", while parliamentary committees debate possible amendments to the PM&DC Act.[2,4,18] ***This has created a paradoxical situation: severe anomalies are acknowledged, but the underlying scoring model remains anchored in unadjusted raw marks.***

Internationally, large-scale admissions tests such as the MCAT, SAT, GRE, GMAT, and LSAT do not rely on raw scores alone when multiple forms and administrations are involved (Table 1). Instead, they use equating and scaled scores to ensure that candidates are not advantaged or disadvantaged by sitting an easier or harder form.[6–11,16,17].

Table 1: Marking Systems in Major International Admissions Exams Versus MDCAT- Pakistan:

EXAM*	PRIMARY REGION	MARKING SYSTEM	PURPOSE
MDCAT	Pakistan	Raw Scores	?
MDCAT	USA/Worldwide	Scaling	Fairness
UCAT	UK/Australia/New Zealand	Scaling	Fairness
SAT	USA, Germany	Scaling	Fairness
CUET	India	Equating	Fairness
GMAT	USA/Canada/World wide	Scaling & Equating	Fairness
LSAT	USA/Canada/World wide	Scaling	Fairness
GRE	Global	Scaling	Fairness

*UCAT: University Clinical Aptitude Test; SAT: Standardized College Admission; CUET: Common University Entrance Test; GMAT: Graduate Management Admission Test; LSAT: Law School Admission Test; GRE: Graduate Record Examination

This article argues that Pakistan's continued reliance on raw scores in a multi-session, multi-year MDCAT is misaligned with international psychometric standards and basic principles of fairness. It also outlines a practical roadmap for reform that is technically feasible and compatible with Pakistan's regulatory context.

METHODS:

This article is based on a narrative, non-systematic review of three central bodies of evidence. First, Pakistani regulatory and legal documents related to MDCAT were examined, including the PM&DC Act 2022 (particularly Section 17 on MDCAT validity) and associated implementation documents and public notices [1,2,12]; MDCAT 2022–2024 information booklets from provincial testing bodies (e.g., University of Health Sciences, UHS) outlining exam structure and score validity [1,12]; and High Court judgments and orders concerning MDCAT 2023–2024, including Sindh High Court decisions referencing score anomalies, paper leaks, and correlation analyses [1,3,4,13–15].

Review Article
https://doi.org/10.63501/hg85j716

Second, media and parliamentary reporting on MDCAT controversies were reviewed, with particular attention to coverage of paper leaks, public protests, and debates around the three-year validity rule and the continued use of raw scores [2,4,18]. Third, international psychometric and testing standards were consulted, including the *Standards for Educational and Psychological Testing* issued by AERA, APA, and NCME [5,6], as well as technical documents and policy statements on equating, scaling, and fairness from organizations such as ETS, NCCAOM, NCBE, and others [7–11,16,17,19,20]. Sources were identified through targeted web searches using terms such as "MDCAT 2024 validity three years," "MDCAT paper leak Sindh High Court," "test equating and scaling," and "standards for educational and psychological testing," and by snowballing from reference lists.

No formal inclusion/exclusion criteria or quality scoring were applied, as the intent was to synthesize policy-relevant evidence rather than estimate effect sizes. Findings are organized into three thematic domains: (1) the limitations of raw scores in a multi-session, multi-year exam; (2) how comparable international exams ensure fairness through equating and scaling; and (3) the ethical implications of current MDCAT practice and a proposed roadmap for reform.

RESULTS:

Raw scores, the simple count or percentage of items answered correctly, are intuitive and acceptable for single-form classroom tests. In a multi-form, multi-session, high-stakes exam like MDCAT, however, raw scores are insufficient. Even with careful blueprinting, no two test forms are exactly equal in difficulty; psychometric studies consistently show that small differences in item difficulty can systematically shift raw-score distributions if they are compared directly.[7–11,16,17] When some candidates receive a slightly easier form and others a slightly harder form, the same underlying ability can yield different raw scores. If these raw scores are then pooled into a single national merit list, candidates are effectively ranked on a mixture of ability and form difficulty. Professional testing standards, therefore, require that scores from different forms or modes be shown to be comparable, generally via equating and scaling procedures.[5,6,10,11]

The fairness risk is magnified when multiple sessions are held across different days and centers, when security breaches or alleged paper leaks affect specific sessions, and when scores are declared valid for multiple years, allowing candidates from different cohorts to compete on unadjusted raw percentages.[1–4,12,18] This is essentially the current MDCAT scenario. Under Section 17(4) of the PM&DC Act 2022, MDCAT scores are valid for three years, and public notices confirm that MDCAT 2023 and 2024 results may be used beyond the year of examination for admissions across Pakistan.[1,2,12] Yet, there is no publicly documented use of form equating or scaled score reporting in MDCAT, and scores are reported simply as percentages out of 200.

MDCAT 2023 and 2024 have been repeatedly challenged in courts and in the public domain. Petitions before the Sindh High Court and other fora have alleged paper leaks, inconsistent invigilation, and irregularities in the compilation of results.[1,3,4,13–15] The Sindh High Court has noted a very weak positive correlation (approximately $r = 0.16$) between MDCAT scores and board examination marks in some datasets,[1] anomalous clusters of candidates scoring 190+ out of 200 (including large groups from specific institutions),[3,4] and sufficient concern to justify stays on result announcements and orders for independent review by investigative and expert bodies.[3,4,13–15]

Media coverage and parliamentary deliberations have described a "validity crisis," with the three-year validity rule and raw-score reporting blamed for creating chaos and perceived unfairness among successive cohorts.[2,4,18] Taken together, these findings point not merely to isolated administrative

lapses, but to structural weaknesses in test design, security, and scoring.

Table 2: Structural Features of MDCAT - Pakistan Compared with Selected International Exams:

Feature	MDCAT (Pakistan)	MCAT (USA)	LSAT (USA)
Primary purpose	Admission to MBBS / BDS	Admission to medical schools	Admission to law schools
Test forms per cycle	Multiple forms across multiple sessions and centers	Multiple equated forms per test window	Multiple equated forms per test window
Reporting metric	Raw score and percentage (out of 200)	Scaled score (472–528)	Scaled score (120–180)
Use of equating	Not publicly documented	Yes, mandatory; detailed technical docs	Yes, mandatory; detailed technical docs
Use of scaling	Not publicly documented	Yes, stable reporting scale	Yes, stable reporting scale
Validity period of scores	Up to 3 years (cross-year use) per PM&DC Act 2022	Typically 2–3 years (school-dependent)	Typically several years (school-dependent)
Treatment of compromised forms	Litigation, court-ordered inquiries, ad hoc decisions	Formal security protocols; potential invalidation and retest	Formal security protocols; potential invalidation and retest
Transparency of methods	Limited public technical reporting	Public technical manuals and validity studies	Public technical manuals and validity statements

By contrast, large-scale admissions and licensure exams internationally treat cross-form and cross-year comparability as a core requirement (Table 2). In these systems, equating is used to statistically link multiple forms so that scores can be interpreted as if they came from the same exam. This typically involves common anchor items and models such as item response theory (IRT), which estimate candidate ability independent of specific items.[7–11,16,17] Scaling then maps equated scores onto a stable reporting scale (for example, 472–528 for the MCAT or 120–180 for the LSAT), so that a score has the same meaning across different administrations.[7,10,11,16,19]

The Standards for Educational and Psychological Testing explicitly state that when multiple forms or modes are used, evidence must be provided that scores are comparable across those forms, including documentation of equating procedures.[5,6] Technical statements from ETS, NCCAOM, NCBE, and others emphasize equating and scaling as essential for ensuring that examinees are not disadvantaged by being assigned more difficult forms.[7–11,16,19,20] ***Pakistan is therefore not being asked to innovate something unprecedented; it is being asked to adopt methods that are already standard practice in the global testing community.***

An admissions test is not a neutral technical instrument; it shapes the composition of the medical workforce for decades. When raw scores from unequated forms are combined and declared valid across multiple years, several ethical consequences follow. First, candidates who sit in easier forms or benefit from compromised security gain a hidden advantage that is not reflected in the score report. Second, students from under-resourced boards, districts, or coaching environments may be further penalized if they receive more difficult forms, compounding existing inequities in the opportunity to learn.[18,21] Third, parents, students, faculty, and the wider public begin to doubt the legitimacy of merit lists and of the regulator itself, with long-term consequences for institutional credibility. ***In a country struggling with maldistribution of doctors and persistent health inequities, basing access to medical education on a statistically fragile and politically contested exam poses profound risks.***

Review Article
https://doi.org/10.63501/hg85j716

DISCUSSION:

Why Raw Scores Are Not Enough

The evidence reviewed here shows that MDCAT operates as a multi-form, multi-session, multi-year high-stakes examination, yet continues to report scores as unadjusted raw percentages. International standards and the technical literature make clear that in such settings, raw scores alone cannot support fair comparisons across forms and years.[5–11,16,17,19,20]

MDCAT 2023-2024 has further revealed how technical shortcomings interact with governance weaknesses. Allegations of paper leaks, weak correlations with board marks, and clusters of very high scores led courts to halt result announcements and order investigations.[1,3,4,13–15] Yet, the regulatory framework has, to date, treated raw scores as final and unchangeable, creating a tension between legal finality and psychometric defensibility.[2,4,18] From an ethical standpoint, continuing to defend raw-score ranking in this environment is challenging to justify.

Towards a Practical Roadmap for Reform

Despite current challenges, Pakistan is fully capable of running an MDCAT that meets international standards of fairness and credibility. A pragmatic reform agenda could include several mutually reinforcing elements (Table 3).

First, equated and scaled scoring should be legally mandated. Parliament and PMDC should explicitly require that any national, high-stakes exam administered in multiple forms or sessions employ psychometric equating and scaled-score reporting, with raw scores used only for internal quality control. If current interpretations of Section 17 of the PM&DC Act are seen as preventing recalculation, they should be amended to allow pre-defined, transparently documented scaling and equating as part of the standard scoring process rather than as an *ad hoc* "normalization".[1,2,12,18].

Table 3: Proposed Roadmap for MDCAT Reform in Pakistan:

Time frame	Action/ Reform Element	Description of Change	Lead/Key Actors
Short Term (≤1–2 years)	Public commitment to move from raw to scaled scores	Announce that from a specified year, reportable scores will be equated and scaled	PMDC; Ministry of Health; provincial authorities
Short Term (≤1–2 years)	Pilot equating on selected forms	Apply anchor-item equating to a subset of MDCAT forms to refine methods	Testing bodies; external psychometric partners
Short Term (≤1–2 years)	Transparency measures	Publish anonymized score distributions and basic item statistics; explain scoring	PMDC; new assessment center (if created)
Medium Term (2–4 years)	Establish National Center for Health Professions Assessment	Create an autonomous body for blueprinting, item development, equating, scaling, and security	Parliament; PMDC; universities
Medium Term (2–4 years)	Implement IRT-based calibration and standard setting	Use IRT models and formal methods (Angoff, Bookmark) for pass/fail and scaling	Assessment center; psychometricians
Medium Term (2–4 years)	Governance and grievance mechanisms	Create independent audit, appeals, and stakeholder panels	Assessment center; civil society
Long Term (≥4 years)	Align MDCAT with competency-based curricula	Coordinate MDCAT content with school curricula; reduce incentives for rote coaching	Education ministries; boards; PMDC
Long Term (≥4 years)	Diversify assessment toolbox	Gradually pilot situational judgment tests, structured interviews, and other tools	Assessment center; medical colleges

Second, an independent National Center for Health Professions Assessment should be created. This autonomous center, analogous to organizations such as NBME or NBOME, would oversee test blueprinting, item development, psychometric analysis, equating, scaling, and ongoing security monitoring. Its governance should include

psychometricians, medical educators, and representatives from public and private universities, with transparent firewalls protecting it from short-term political pressure.[7–11,16,17,19,20]

Third, modern test-development and equating methods should be adopted. This would involve designing forms with common anchor items to enable robust equating; implementing IRT-based calibration to ensure candidate ability estimates are comparable across forms; and using established standard-setting methods (such as Angoff or Bookmark) to set pass marks rather than *ad hoc* percentage cut-offs.[7–11,16,17,19,20] Capacity-building partnerships with international psychometric centers could accelerate this transition.

Fourth, transparency and public communication need to be enhanced. Technical improvements alone will not restore trust. PMDC and any new assessment center should publish annual technical reports, provide plain-language explanations of scaled scores, and establish independent grievance and audit mechanisms with representation from students, parents, and faculty.[18,21]

Fifth, MDCAT should be aligned with broader educational goals. Over-reliance on a single multiple-choice exam encourages rote learning and a commercial coaching culture. MDCAT reform should therefore be coordinated with efforts to strengthen competency-based school curricula and, over time, to pilot additional tools such as situational judgment tests and structured interviews for assessing non-cognitive attributes relevant to medical practice.[17,21]

Short-Term Steps for MDCAT 2025 and Beyond

Even before full structural reform is achieved, several immediate steps are feasible. From MDCAT 2026 onward, regulators could publicly commit to reportable scores based on equated and scaled metrics rather than raw percentages, with independent psychometric oversight. In parallel, equating procedures could be piloted on a subset of forms in 2026 to refine methods prior to full implementation.

Regulators should also release anonymized item statistics and score distributions by region to enable external expert review and actively engage medical universities so that they understand and support the use of scaled scores rather than attempting to reverse-engineer them into raw percentages. Taken together, such steps would signal that Pakistan is serious about fairness at the entry gate to medical education.

CONCLUSION:

The MDCAT crisis is not solely about leaked papers, anomalous score clusters, or disputed merit lists; it is a mirror reflecting deeper weaknesses in governance, technical capacity, and accountability in Pakistan's assessment system. When a high-stakes national exam is administered in multiple forms, across multiple sessions and years, yet continues to rely on unadjusted raw percentages, it exposes a structural misalignment between the complexity of the task and the tools being used to accomplish it. Courts, the media, and the public have all been drawn into disputes that, in a well-designed system, would have been anticipated and prevented by robust psychometric procedures and transparent oversight.

Pakistan now faces a clear choice. One path is to continue defending a raw-score system that is psychometrically indefensible in a multi-session, multi-year environment, thereby normalizing *ad hoc* crisis management, repeated litigation, and the erosion of public trust. The alternative is to seize this moment as an inflection point: to legislate for equated and scaled scoring, establish an independent national assessment center, and embed modern test development, security, and reporting practices into the fabric of MDCAT. This is not a question of uncritically copying foreign models, but of aligning a national exam with sound measurement principles already standard across reputable admissions and licensure tests worldwide.

Scaling and equating are not abstract statistical luxuries reserved for wealthy countries; they are instruments of justice and fairness. In a context where educational opportunity is uneven and access to competent health care can be literally life-saving, fairness at the first gate that decides who may train as a doctor or dentist becomes an ethical imperative rather than a technical detail. When unequated raw scores from different forms and years are treated as interchangeable, invisible advantages and disadvantages are baked into the merit list, compounding existing inequities by geography, schooling, and socioeconomic status. ***Reforming MDCAT scoring is therefore not only a psychometric necessity but also a matter of distributive justice and professional integrity.***

If Pakistan can reform MDCAT along the lines proposed in this article, mandating equating and scaling, creating an autonomous assessment authority, strengthening test security, and communicating transparently, it will do more than produce credible score reports. It will begin to renew the social contract between aspiring students, their families, and the institutions that promise to train the next generation of healers. Such a system would signal to the public that merit is being measured carefully and fairly, that regulators are willing to be guided by evidence and global best practice, and that the privilege of entering the medical and dental professions rests on a foundation that is both technically sound and morally defensible.

Use of Generative Artificial Intelligence:

Generative artificial intelligence (Claude, Anthropic) was used solely to assist with language refinement and copy-editing; all ideas, arguments, and conclusions are those of the author, who takes full responsibility for the content.

REFERENCES

1. **High Court of Sindh.** Constitutional petitions regarding MDCAT-2024 (judgment noting correlation between MDCAT and board marks). 2024 Oct 26. https://caselaw.shc.gov.pk/caselaw/view-file/MjIxNTUzY2Ztcy1kYzgz.
2. **Pakistan Medical and Dental Council (PMDC).** Public Notice regarding MDCAT-2024. Islamabad: PMDC; 2024. https://pmdc.pk/home/PublicAnnouncement (see "Public Notice Regarding MDCAT-2024").
3. **High Court of Sindh.** Order staying MDCAT-2024 result announcement and directing investigation into alleged paper leaks. 2024 Oct 10. Official order accessible via the Sindh High Court case-law database (https://caselaw.shc.gov.pk – search "MDCAT-2024" orders dated 10 Oct 2024). https://www.thenews.com.pk/print/1238588-shc-stays-announcement-of-mdcat-results.
4. **Dawn.** SHC stays MDCAT results, orders probe into paper leak. *Dawn.* 2024 Oct 10. https://www.dawn.com/news/1864127.
5. **American Educational Research Association, American Psychological Association, National Council on Measurement in Education.** *Standards for Educational and Psychological Testing.* Washington, DC: AERA; 2014. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf.
6. **American Psychological Association.** The Standards for Educational and Psychological Testing. https://www.apa.org/science/programs/testing/standards.
7. **Angoff WH.** *Scales, Norms, and Equivalent Scores.* Princeton, NJ: Educational Testing Service; 1984. https://www.ets.org/research/policy_research_reports/publications/book/1984/ibmt.
8. **National Certification Commission for Acupuncture and Oriental Medicine (NCCAOM).** *Equating and Scaling: Assuring the Highest Level of Fairness to Candidates.* 2017. https://www.nccaom.org/wp-content/uploads/pdf/Document%20B%20Equating%20and%20Scaling.pdf.
9. **Winter PC.** *Evaluating the Comparability of Scores from Achievement Test Forms.*

Washington, DC: Council of Chief State School Officers; 2010.
10. **Bellwether Education Partners.** *Demystifying Statewide Standardized Assessments: Ensuring Comparability Across Administrations.* Washington, DC: Bellwether; 2023. https://bellwethereducation.org/resource/demystifying-statewide-standardized-assessments-ensuring-comparability-across-administrations.
11. **Assessment Systems Corporation.** Test score equating and linking. Assessment Systems blog; c2020–2023. https://assess.com/test-equating/.
12. **University of Health Sciences (UHS).** *MDCAT Information Booklet 2023.* Lahore: UHS; 2023. https://www.uhs.edu.pk/mdcatadmissionnews.php (see "MDCAT Information Booklet 2023").
13. **The Express Tribune.** SHC issues notices against re-leak of MDCAT paper. *The Express Tribune.* 2023 Dec 5. https://tribune.com.pk/story/2448946/shc-issues-notices-against-re-leak-of-mdcat-paper.
14. **Dawn.** Allegations of MDCAT paper leak surface again as students move to SHC. *Dawn.* 2023 Dec 5. https://epaper.dawn.com/DetailNews.php?StoryText=05_12_2023_113_001.
15. **Pakistan Medical and Dental Council (PMDC).** Public announcements on the validity of the National MDCAT 2022 result and the MDCAT 2023–2024 policy. Islamabad: PMDC; 2023–2024. https://pmdc.pk/home/PublicAnnouncement.
16. **Wyse AE.** Two statistics for measuring the score comparability of computerized adaptive tests. *Applied Psychological Measurement.* 2023;47(7–8):513–525. doi:10.1177/01466216231209749. https://pubmed.ncbi.nlm.nih.gov/37997578/ and https://journals.sagepub.com/doi/10.1177/01466216231209749
17. **Meagher DG.** *Test Bias, Fairness, and Standardized Admission Tests.* Bloomington, MN: Pearson; 2020. https://www.aacp.org/sites/default/files/2020-09/test-bias-fairness-and-standardized-admission-tests.pdf.
18. **Minute Mirror.** Three-year validity rule sparks MDCAT chaos; Parliament moves to amend PMDC Act. *Minute Mirror.* 2025 Nov 20. https://minutemirror.com.pk/three-year-validity-rule-sparks-mdcat-chaos-parliament-moves-to-amend-pmdc-act-464769/.
19. **Educational Testing Service.** Fairness and validity of the GRE tests. Princeton, NJ: ETS; c2019–2024. https://www.ets.org/gre/score-users/about/fairness.html
20. **National Conference of Bar Examiners (NCBE).** The Testing Column: Assessment scales and fairness in assessment. *The Bar Examiner.* 2021–2023.