

Generation of Medical Case-Based Multiple-Choice Questions

Somaiya Al Shuriaqi ^{1,*}, Abdulrahman Aal Abdulsalam ¹ and Ken Masters ²

¹ Department of Computer Science, College of Science, Sultan Qaboos University, P.O. Box 243, Muscat 123, Oman; a.aalabdulsalam@squ.edu.om

² Medical Education and Informatics Department, College of Medicine and Health Sciences, Sultan Qaboos University, P.O. Box 243, Muscat 123, Oman; kmasters@squ.edu.om

* Correspondence: somaiya@squ.edu.om

Abstract: This narrative review is a detailed look at how we make multiple-choice questions (MCQs) based on medical cases in today's medical teaching. Moving from old-style MCQs to ones that are more related to real clinical situations is really important. It helps in growing critical thinking and practical use, especially since MCQs are still the primary method for testing knowledge in medicine. We look at the history, design ideas, and both manual and computer-based methods that have helped create MCQs. Technologies like Artificial Intelligence (AI) and Natural Language Processing (NLP) are receiving a lot of focus for their ability to automate the creation of question. We also talk about the challenges of using real patient cases, like the need for exact clinical information, reducing unclear information, and thinking about ethical issues. We also investigate the measures of validity and reliability that are crucial to maintaining the honesty of case-based MCQs. Finally, we look ahead, speculating on where medical education is headed as new technologies are incorporated and the value of case-based evaluations continues to rise.

Keywords: medical case-based multiple-choice questions (CB-MCQs); distractors; artificial intelligence (AI); natural language processing (NLP)



Citation: Al Shuriaqi, S.; Aal Abdulsalam, A.; Masters, K. Generation of Medical Case-Based Multiple-Choice Questions. *Int. Med. Educ.* **2024**, *3*, 12–22. <https://doi.org/10.3390/ime3010002>

Academic Editor: Hideki Kasuya

Received: 1 November 2023

Revised: 20 December 2023

Accepted: 21 December 2023

Published: 25 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The field of medical education has consistently been at the cutting edge of integrating innovative approaches and strategies to improve student learning and evaluation due to its dynamic character. The multiple-choice question (MCQ) is a significant evaluation tool that has been well recognized for its efficiency, objectivity, and capacity to encompass a wide range of knowledge in a brief manner [1]. MCQs were initially implemented with the intention of optimizing the testing procedure and establishing uniformity. However, they have since progressed to encompass broader educational goals, covering not only memorization but also the development of critical thinking skills and the ability to make clinical decisions [2].

Nevertheless, due to the growing emphasis on the practicality of clinical knowledge and the significance of problem-solving abilities in the medical education framework, there has been a shift in the approach towards including case-based MCQs. These kinds of questions offer students the opportunity to engage with clinical scenarios that they may experience during their professional practice, therefore facilitating the integration of theoretical knowledge with practical implementation [3]. The aforementioned transformation is not solely a result of pedagogical inclination, but rather stems from the necessity for physicians to possess proficiency in navigating practical clinical predicaments and in rendering well-informed judgments.

The rapid expansion of technology, particularly the emergence of Artificial Intelligence (AI) and Natural Language Processing (NLP), has significantly enhanced this field, providing resources that can assist in the automated creation of case-based MCQs. The previously

mentioned developments exhibit potential in enhancing the process of question production, guaranteeing a broader scope, and delivering personalized learning experiences [4].

The purpose of this narrative review is to provide a thorough examination of the development of medical case-based MCQs, covering their origins and recent technology breakthroughs and discussing their importance, methodology, and potential future directions [5–8].

1.1. Historical Background

MCQs were first developed during the early 20th century, representing a significant departure from the conventional essay-based exams that were commonly used during that era [9]. MCQs are appealing, standardized assessments which have the ability to bring uniformity to the evaluation process, hence guaranteeing a consistent gauge of students' knowledge and skills within extensive groups. According to Stough (1993), this particular structure enabled the use of objective grading methods and streamlined the evaluation process for a diverse range of subjects, all within a constrained timeframe [6].

1.2. Structure of a Case-Based MCQ

Irrespective of their role, case-based MCQs follow a standard format [10,11] as follows:

- The stem (sometimes the portion referred to as the “question”). This might consist of a simple question, but might also be more complex, and include a scenario and media. The key element in creating a robust multiple-choice question is to ensure that the stem is well-defined and focused. The stem of the query must contain the primary concept.
- Alternatives (sometimes referred to as “options”). These include all the items, from which the user must select one.
- Answer (sometimes referred to as the “correct answer” or the “key”). This is one of the alternatives, and is the actual required answer to the question. The crucial characteristic is that the selected option deemed as accurate must be absolutely indisputable, without any doubt or debate. It is preferable to have a manuscript citation or reference on hand for verification purposes. When providing a correct answer, beware of using ambiguous phrases like “frequently”, “often”, “rarely”, or “sometimes”. These hints indicate that an answer is correct and demonstrate test-taking intelligence rather than subject content knowledge.
- Distractors. There are all the alternatives that are not the answer. From a cognitive perspective, it is acceptable to have two distractors. However, in health sciences testing, it is more common to have three or four distractors. Writing plausible distractors can be the most difficult aspect of developing a well-formulated examination.

Table 1 gives an example showing the constituents of a case-based MCQ [4].

Table 1. The constituents of a case-based MCQ. Note that the Answer (key) and the Distractors together form the Alternatives.

Case-Based MCQ Example	Constituent
A 50-year-old man has had gradually progressive hand weakness. He has atrophy of the forearm muscles, fasciculations of the muscles of the chest and arms, hyperreflexia of the lower extremities, and extensor plantar reflexes. Sensation is not impaired. Which of the following is the most likely diagnosis?	Stem
A. Amyotrophic lateral sclerosis	Answer (or key)
B. Dementia, Alzheimer’s type	Distractor 1
C. Guillain–Barré syndrome	Distractor 2
D. Multiple cerebral infarcts	Distractor 3
E. Multiple sclerosis	Distractor 4

1.3. Transition to Clinical Significance

Initial MCQs primarily emphasized the retrieval of factual information. However, educators swiftly acknowledged the necessity of assessing more advanced cognitive abilities, particularly within the intricate and diverse field of medicine. During the 1980s, there was an increasing focus on the alignment of MCQs with clinical scenarios, thereby replicating authentic medical situations that students may face during medical practice [12]. The transition discussed in this context was motivated by a pedagogical shift towards problem-based and team-based learning. This approach placed greater importance on the application of acquired knowledge in clinical settings, as opposed to solely focusing on knowledge acquisition [13].

1.4. Emergence of Case-Based MCQs

The advent and fast acceptance of case-based MCQs in the late 20th and early 21st centuries might be seen as a continuation of the focus on clinical relevance. The questions presented in this study were based on actual or simulated patient scenarios, and examinees were expected to apply their knowledge, analyze clinical data, and make well-informed decisions. These activities closely resembled the responsibilities of a practicing physician [4].

1.5. Integration of Technology

The emergence of the digital era brought about a significant transformation in the development and administration of MCQs. The prevalence of computer-based testing has led to the emergence of increasingly interactive and dynamic question styles. Concurrently, the incorporation of Artificial Intelligence and data analytics caused an impact on the construction of MCQs, presenting the possibility of customization and adaptive testing [2].

It can be inferred that the aforementioned points collectively support the notion that MCQs in medical education have evolved in parallel with the broader educational and technological advancements within the profession, progressing from their modest origins to their present complex forms. The enduring significance of their contribution to the development of proficient and analytically minded medical practitioners is unquestioned [14,15].

2. The Significance of Case-Based MCQs

The primary objective of medical education is not only to provide students with fundamental knowledge, but also to provide them with the abilities required to effectively use this knowledge in practical clinical situations. Case-based MCQs are an essential tool in this pursuit, providing numerous unique benefits.

2.1. The Integration of Theory and Practice

Case-based MCQs provide a connection between theoretical medical principles and practical clinical scenarios. The questions presented to students involve real or theoretical patient scenarios, which necessitate the navigation of complicated clinical reasoning. This approach aims to foster a more profound comprehension and practical application of academic information [14].

2.2. Evaluating Higher-Order Cognitive Abilities

Traditional MCQs frequently assess the ability to recall factual information. On the other hand, case-based MCQs require the utilization of advanced cognitive abilities, such as analysis, application, and evaluation. The promotion of critical thinking and decision-making skills, which are fundamental abilities for healthcare professionals, is achieved by involving the students in clinical vignettes [16].

2.3. Improving Clinical Readiness

The process of clinical decision-making encompasses more than the simple recollection of knowledge; it necessitates the integration of information within the limitations

of ambiguity and time sensitivity. According to Chéron et al., case-based MCQs effectively replicate these obstacles, hence enhancing students' readiness for real-world clinical responsibilities [5].

2.4. Embracing Contemporary Pedagogical Approaches

The trend towards problem-oriented combined learning in medical education is well-supported by the utilization of case-based MCQs. Zhao et al. posits that student-centered education is promoted by fostering an environment that encourages students to actively engage in the process of learning [15].

2.5. The Provision of Objective Assessment Metrics

Although case-based MCQs contain a substantial amount of information, they continue to possess the inherent objectivity associated with the MCQ format. Zhao et al. assert that the implementation of this approach guarantees impartial evaluation and provides quantifiable measures that may be utilized for the goals of feedback, enhancing the curriculum, and achieving accreditation [15].

2.6. Enhancing Proficiency in Differential Diagnosis Abilities

Frequently, case-based MCQs pose scenarios wherein symptoms may correspond to many illnesses, necessitating students to discern and rank potential diagnoses based on their relative importance. Engaging in this activity enhances their proficiency in differential diagnosis, a fundamental component of clinical practice [14].

Case-based MCQs serve a dual purpose beyond mere evaluation, as they play a crucial role in developing a prospective physician's clinical expertise by facilitating the integration of theoretical knowledge with practical application. These questions effectively bridge the divide between academic learning and real-world medical practice. The role they play in contemporary medical education is undeniably essential and significant [6].

3. Approaches for Generating Case-Based MCQs

The creation of case-based MCQs, which play a crucial role in evaluating practical medical knowledge, can be accomplished using both conventional human techniques and novel automated methods. Every methodology inherently possesses its unique array of benefits and challenges.

3.1. Generation through Manual Procedures

In the ensuing subsections, a detailed exploration regarding the manual generation procedure will be undertaken. Manual generation delineates a methodology whereby textual or content creation is executed solely by human authors, devoid of any incorporation or interference from technological apparatuses or systems. The manual formulation of case-based MCQs occurs through several stages:

- a. **Selection of Topic:** During the initial phase, educators execute a meticulous selection of a medical subject or issue that bears relevance to the curriculum, as indicated by Al-Rukban [7]. If the institution uses Learning Objectives, then these must also be noted to ensure that the questions are aligned with them.
- b. **Development of Case Scenario:** This phase entails the crafting of a patient scenario which could be derived from either authentic experiences or hypothetical situations, aiming to construct a contextual framework. Typical elements of a patient's medical record integrate their medical history, vital statistics, laboratory results, and other pertinent information [14].
- c. **Question Framing:** The core objective of framing questions is to evaluate understanding, analysis, or application in connection with the presented case study [7,14].
- d. **Distractor generation:** Distractor conceptualization involves formulating conceivable incorrect alternatives (distractors) that are coherent and non-deceptive, a notion underscored by Al-Rukban and Kurdi [7,14].

- e. **Validation:** The refinement and validation of MCQs are optimized through a peer review, executed by educationalists and clinical experts. This collaborative methodology ascertains the enhancement of question clarity, accuracy, and pertinence [14].

3.2. Challenges of Manual Generation

There are several challenges and constraints associated with the generation of case-based MCQs. The process of creating case-based MCQs of superior quality using manual means can be a time-consuming endeavor, as noted by Leo et al., (2019) [4]. Also, the presence of prejudice is a possibility in educational settings, as the personal biases held by educators might potentially impact the formulation and presentation of questions. In addition, the diversity of manually constructed MCQs may be limited, as they may not cover the full range of probable clinical circumstances or question styles [17].

3.3. The Process of Automated Generation

The advent of digital transformation in the field of education has given rise to the utilization of artificial intelligence (AI) and natural language processing (NLP) as highly effective instruments. According to Zhang et al. these systems possess the capability to analyze extensive quantities of text, detect patterns within the data, and provide queries that are contextually appropriate [18].

The tools and techniques are employed in the process of generating MCQs automatically. Firstly, the database-driven approach involves the utilization of algorithms to extract information from medical ontology or texts in order to generate questions that are grounded in current, evidence-based content [6]. In addition, Natural Language Processing (NLP) tools such as Generative Pre-trained Transformer (GPT) and Bidirectional Encoder Representations from Transformers (BERT) are utilized to analyze medical texts. These tools are capable of extracting essential concepts and relationships from the texts, enabling the generation of preliminary MCQs [19]. Additionally, according to Torrealba et al., adaptive learning systems, which are powered by artificial intelligence, have the capability to modify the difficulty level of MCQs based on the performance of individual students [19]. This adaptive approach aims to optimize the learning process.

3.4. The Advantages and Obstacles Associated with Automated Generation

There exist numerous advantages associated with the automated creation of case-based MCQs. Firstly, the ability to generate a substantial quantity of questions quickly is considered an important aspect of efficiency in question creation [4]. In addition, the continuous updates from medical databases guarantee the relevance of the content. Also, diversity can be achieved by incorporating a range of question styles and clinical settings.

One potential issue with automatically generated MCQs is the potential lack of depth and clinical relevance in the absence of human scrutiny [6]. Also, the potential degradation of the educator's role in curriculum design is a significant ethical consideration associated with the over-reliance on artificial intelligence (AI).

4. Principles for Designing Case-Based MCQs

The creation of efficient MCQs that are based on specific cases requires the integration of a comprehensive knowledge of the subject matter and the use of robust design guidelines. These guidelines ensure that MCQs not only evaluate knowledge but also align with the intended educational and assessment goals.

4.1. Authenticity of Cases

A key concept in the formation of case-based MCQs involves the crucial aspect of verifying the validity of the scenarios. It is crucial that these scenarios accurately resemble authentic clinical circumstances that clinicians commonly experience. This approach is based on the premise that incorporating real-life scenarios can improve students' capacity to apply their theoretical knowledge in practical clinical settings. These exemplars act as

instrumental methods to equip students with the essential capabilities and insights needed to adeptly navigate prospective practical impediments encountered in their forthcoming professional pursuits [17].

4.2. Precision and Clarity in Language

The imperative of clarity cannot be overstated in formulating scenarios and the subsequent inquiries, necessitating the employment of precise and unambiguous language whilst omitting superfluous details. A meticulously composed MCQ ensures the assessment of students on accurate medical knowledge and reasoning capabilities, rather than their prowess in interpreting intricate or obscure language. The cornerstone of impartial assessment lies in clear articulation, ensuring equal accessibility for all examinees [20].

4.3. Pertinent and Rigorous Distractors

Crafting exemplary distractors demands an amalgamation of artistic and analytical aptitudes. Distractors need to be plausible and closely connected to the scenario provided to ensure a rigorous selection process. The aim is to formulate distractors that can discriminate between students with genuine subject mastery and those with superficial or misconstrued understanding. Effective distractors, as per Salam et al., can provide profound insights into students' understanding levels [21].

4.4. Incorporation of Clinical Reasoning

Beyond assessing rudimentary factual knowledge, case-based MCQs are robust indicators of clinical reasoning capabilities. It is crucial for educators to foster critical analysis, interpretation, and the pragmatic application of accumulated knowledge among students. By integrating clinical reasoning within assessments, educators can ensure that students are assimilating and applying knowledge adeptly within medical paradigms [22].

4.5. Facilitation of Feedback

Learning is an ongoing process that extends beyond answering; constructive feedback catalyzes and refines learning experiences. Providing explanations for both correct and incorrect responses is recommended, acting as an efficacious mechanism for solidifying knowledge, correcting misinterpretations, and deepening understanding [23].

4.6. Conformity with Educational Objectives

Every MCQ should align with the educational goals and aims of the curriculum or specific module being addressed. The alignment of assessment tools is crucial in ensuring their relevance and fairness, since they accurately measure the specific knowledge and skills that students are expected to acquire [24].

4.7. Recurrent Evaluation and Validation

Given the evolving landscape of medicine, punctuated by regular breakthroughs and innovations, continual reassessments and validations of MCQs are essential to uphold their accuracy and relevance. Employing an ongoing review and validation strategy assures that assessments remain congruent with the continual transformations within the medical domain [25].

The intricate composition of case-based MCQs necessitates an intricate balance of content expertise and foundational educational principles to adequately address and reflect both the evolving nature of medical knowledge and the diverse learning needs of students. When formulated in accordance with these guiding principles, MCQs prove to be important instruments for both instructional purposes and assessment, playing a crucial role in shaping the future cohort of medical professionals.

5. Validity and Reliability of Case-Based MCQs

The effectiveness of any assessment instrument, such as case-based MCQs, is predicated on two essential factors: validity and reliability. In order to fulfill their intended objectives with efficacy, it is imperative that these MCQs yield precise and reliable outcomes across diverse situations and populations.

5.1. Validity of Case-Based MCQs

Validity describes how well a study or a test really measures what it is supposed to measure. It is a key idea in testing and crucial in the context of case-based MCQs. These MCQs should test whether students can use their medical knowledge in real-life situations. When we talk about the validity of case-based MCQs, there are three main types that can be considered:

- **Content validity:** This measures the extent to which the MCQs cover the topics and relevant clinical areas. It is important to select cases that show the many different situations that doctors might face [6].
- **Criterion-related validity:** This measures the extent to which the results from these MCQs match results from other tests measuring the same skills or knowledge. For example, you could check how well scores from a case-based MCQ test compare with scores from a hands-on clinical exam, following the ideas of Messick et al. [26]. Criterion-related validity can be further divided into concurrent validity (comparing to an established tool) and predictive validity (for predicting future outcomes) [27].
- **Construct validity:** This measures the extent to which the MCQs test the theoretical ideas that are supposed to be assessed. The questions should be able to effectively check if the intended theoretical concepts are understood by the students. For example, if a collection of MCQs are formulated with the intention of evaluating clinical reasoning abilities, it may be hypothesized that students who achieve higher scores on these assessments will possess superior clinical reasoning skills compared to those who obtain lower scores [28].

5.2. Reliability of Case-Based MCQs

The concept of reliability measures the degree of consistency exhibited by assessment outcomes. A case-based MCQ that is reliable would demonstrate consistent outcomes when administered to the same student on multiple occasions or when assessed by different examiners.

- **Test-retest reliability:** Test-retest reliability refers to the assessment of the consistency of scores obtained by students when they take the same test on several occasions. According to Nunnally, a strong correlation between the MCQs indicates that they yield consistent outcomes throughout different time periods [29].
- **Internal consistency:** Internal consistency refers to the degree to which several items within an MCQ test produce consistent outcomes. Cronbach's alpha is a frequently employed statistical measure for this particular objective. According to Nunnally, a high result, often over 0.7, signifies that the MCQs effectively assess the same underlying construct [29]. For test reliability, however, the Kuder-Richarson 20 (KR20) (when item difficulty is variable) and the KR21 (when the item difficulty is similar) are preferred [30].
- **Inter-rater reliability:** Inter-rater reliability is of utmost importance in situations where questions are open-ended and necessitate manual scoring. The measure assesses the level of consensus among several raters or examiners. According to Kurdi et al., a high coefficient of inter-rater reliability signifies a consistent scoring pattern among many examiners [6].

To be considered effective, case-based MCQs must adhere to the stringent criteria of both validity and reliability. The utilization of regular evaluations, in conjunction with

statistical analysis, can assist educators in enhancing MCQs, ensuring their continued reliability and validity as instruments for evaluating clinical knowledge and reasoning.

Although not directly related to validity and reliability, one can also consider other measures to assist in evaluating an MCQ test. Among these is a more detailed item analysis that can be used to check the level of difficulty of each test item and the discrimination index (to measure the extent to which this item differentiates between weaker and stronger students). In most instances, if the test is delivered online, then the test-delivery software calculates these automatically. For more on these, the reader can refer to Carneson et al. [10].

6. Challenges and Controversies in Case-Based MCQs

Case-based MCQs have played an essential part in medical education, providing a controlled method to evaluate clinical reasoning. Nevertheless, their use is not devoid of difficulties and controversy. Numerous challenges are involved with medical case-based MCQs.

- **Overemphasis on Recall:** One potential issue with MCQs is the tendency to place excessive emphasis on recall. Although MCQs are effective in evaluating a wide range of content, there is a valid concern that they primarily measure memorization rather than comprehensive comprehension or practical application [3]. In order for case-based MCQs to be truly effective, it is imperative that they redirect their emphasis from mere memory to the domains of application and synthesis.
- **The Importance of Misleading Distractors:** The inclusion of effective distractors plays a vital role in enhancing the discriminatory capacity of MCQs. Nevertheless, it is worth noting that inadequately formulated distractors have the potential to deceive students, transforming questions into assessments of test-taking abilities rather than evaluating their clinical reasoning capabilities [31].
- **Dependence on Stem Clarity:** The importance of stem clarity in determining dependence is of utmost significance. The presence of ambiguous or unnecessarily lengthy stems in academic assessments may lead to unanticipated challenges, which could potentially put learners at a disadvantage [31].
- **Cultural and Socio-Economic Biases:** MCQs may unintentionally include cultural or socio-economic biases, thereby reflecting the perspectives of the individuals who authored the questions rather than universally accepted medical information. Kim and Zabelina (2015) argue that prejudices have the potential to create disadvantages for specific student groups [32].
- **Over-Reliance on Single Best Answer:** The practical situations encountered in the field of medicine rarely possess unequivocal resolutions. According to Scott et al. (2018), the utilization of single-best-answer MCQs may occasionally result in an oversimplification of intricate clinical settings [8].
- **Security Concerns:** Security concerns have emerged due to the widespread use of digital platforms and student collaboration tools, which have raised apprehensions over the security of exams and the potential for question-sharing. These concerns pose significant risks to the overall integrity of the assessment process [33].
- **Technology Dependence:** The integration of MCQs into digital platforms has led to a heightened reliance on technology. This presents difficulties pertaining to software malfunctions, user interface usability, and the issue of digital equity [34].

It is imperative to recognize and confront the issues associated with case-based MCQs in medical education in order to maintain their effectiveness and impartiality as assessment tools.

7. Future Directions

The field of medical education and evaluation is constantly changing, with ongoing developments in the design and utilization of case-based MCQs offering potential for creative breakthroughs.

According to Larranaga et al., there is a growing trend towards a greater incorporation of Artificial Intelligence (AI) and Natural Language Processing (NLP) in the development and improvement of case-based MCQs [35]. This combination holds the potential to enhance learning experiences by providing more personalized and adaptive approaches. The recent progress in these technologies will facilitate the creation of more intricate evaluation tools capable of conducting a comprehensive analysis of students' replies and offering valuable feedback.

The integration of Virtual and Augmented Reality (VR/AR) technology has the potential to enhance the intricacy and authenticity of case scenarios, facilitating immersive learning experiences [36]. These advances have the potential to improve the capacity of MCQs to evaluate higher-order cognitive abilities and clinical reasoning within simulated clinical settings.

The use of more advanced security measures is anticipated in the future to address the issue of cheating and uphold the authenticity of online assessments [37]. The maintenance of online examination validity necessitates the implementation of safe browsing technologies and innovative testing systems.

The forthcoming trend is expected to place greater importance on the recognition and mitigation of cultural and socio-economic biases in MCQ construction, with the aim of promoting inclusion and equality in assessment practices [32]. The integration of expertise from subject matter professionals and diversity specialists can play a significant role in the development of impartial and fair assessment instruments.

Ongoing verification processes are essential to guarantee that case-based MCQs stay relevant and compatible with the growing medical educational programs. These ongoing developments necessitate the establishment of periodic review and validation processes. This is in accordance with the guidelines provided by the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) in 2014 [34].

The integration of adaptive learning technologies and analytical tools will facilitate the development of customized learning paths, which will align assessments with the unique requirements and progress of learners [38].

8. Conclusions

The exploration into the development of medical case-based MCQs has spanned its historical evolution, with foundational design principles, and advancements in methodology. The aim of this review was to provide an exhaustive analysis of the intricate processes involved in formulating medical case-based MCQs, concentrating particularly on their pivotal role in assessing clinical reasoning and decision-making competencies within medical education.

Inherent principles governing the construction of case-based MCQs encapsulate a pronounced emphasis on content validity, reliability, and the amalgamation of clinical reasoning. These principles fortify the reliability and credibility of MCQs as an assessment instrument, amplifying their efficacy in orienting medical students for their impending responsibilities by harmonizing them with pragmatic clinical environments.

However, as underscored in the segment addressing challenges and controversies, the deployment of case-based MCQs is not exempt from its inherent limitations and disadvantages. Ongoing attention and resolutions are required to uphold the reliability and fairness of assessments due to several concerns, including cultural and socio-economic biases, an excessive focus on recall, and security considerations.

The potential of case-based MCQs in the future appears to be abundant, as technological breakthroughs such as artificial intelligence (AI), natural language processing (NLP), and virtual reality/augmented reality (VR/AR) hold promise in improving the overall quality, security, and adaptability of these assessments. These technologies have the potential to provide enhanced and individualized learning experiences, thereby playing a crucial role in catering to the varied learning requirements of medical students.

Moreover, the growing recognition and endeavors aimed at tackling diversity, inclusion, and equity in medical education indicate the need for a comprehensive strategy for developing case-based MCQs in the coming years. The implementation of continuing measures aimed at mitigating biases and guaranteeing the pertinence and impartiality of evaluations represents essential actions in the pursuit of educational equity.

In summary, the creation of medical case-based MCQs is a complex and dynamic area, requiring ongoing improvements and adjustments to align with the growing contexts of medical education and technology. Although the process is accompanied by various difficulties, the persistent efforts to surmount these obstacles and introduce novel approaches are influencing a future that strives to attain enriched educational experiences, fair evaluations, and finally enhanced healthcare for patients.

Author Contributions: Writing—original draft preparation, S.A.S.; review and editing, A.A.A. and K.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Mujeeb, A.; Ghongane, B.; Pardeshi, M. Comparative assessment of multiple choice questions versus short essay questions in pharmacology examinations. *Indian J. Med. Sci.* **2010**, *64*, 118. [[CrossRef](#)] [[PubMed](#)]
- Bassett, M.H. Teaching Critical Thinking without (Much) Writing: Multiple-Choice and Metacognition. *Teach. Theol. Relig.* **2016**, *19*, 20–40. [[CrossRef](#)]
- Khan, M.-U.; Aljarallah, B.M. Evaluation of Modified Essay Questions (MEQ) and Multiple Choice Questions (MCQ) as a tool for Assessing the Cognitive Skills of Undergraduate Medical Students. *Int. J. Health Sci.* **2011**, *5*, 39–43.
- Leo, J.; Kurdi, G.; Matentzoglou, N.; Parsia, B.; Sattler, U.; Forge, S.; Donato, G.; Dowling, W. Ontology-Based Generation of Medical, Multi-term MCQs. *Int. J. Artif. Intell. Educ.* **2019**, *29*, 145–188. [[CrossRef](#)]
- Chéron, M.; Ademi, M.; Kraft, F.; Löffler-Stastka, H. Case-based learning and multiple choice questioning methods favored by students. *BMC Med. Educ.* **2016**, *16*, 41. [[CrossRef](#)] [[PubMed](#)]
- Kurdi, G.; Leo, J.; Parsia, B.; Sattler, U.; Al-Emari, S. A Systematic Review of Automatic Question Generation for Educational Purposes. *Int. J. Artif. Intell. Educ.* **2019**, *30*, 121–204. [[CrossRef](#)]
- Al-Rukban, M. Guidelines for the construction of multiple choice questions tests. *J. Fam. Community Med.* **2006**, *13*, 125–133. [[CrossRef](#)]
- Scott, K.R.; King, A.M.; Estes, M.K.; Conlon, L.W.; Phillips, A.W. Evaluation of an Intervention to Improve Quality of Single-best Answer Multiple-choice Questions. *WestJEM* **2018**, *20*, 11–14. [[CrossRef](#)]
- Stough, L.M. Research on Multiple-Choice Questions: Implications for Strategy Instruction. In *Annual Convention of the Council for Exceptional Children*, 71st ed.; Council for Exceptional Children: San Antonio, TX, USA, 1993; pp. 1–11.
- Carneson, J.; Delpierre, G.; Masters, K. *Designing and Managing Multiple Choice Questions*, 2nd ed.; Cape Town, South Africa, 2016. [[CrossRef](#)]
- DiSantis, D.J. A Step-By-Step Approach for Creating Good Multiple-Choice Questions. *Can. Assoc. Radiol. J.* **2020**, *71*, 131–133. [[CrossRef](#)]
- Vuma, S.; Sa, B. A comparison of clinical-scenario (case cluster) versus stand-alone multiple choice questions in a problem-based learning environment in undergraduate medicine. *J. Taibah Univ. Med. Sci.* **2016**, *12*, 14–26. [[CrossRef](#)]
- Stringer, J.K.; Santen, S.A.; Lee, E.; Rawls, M.; Bailey, J.; Richards, A.; Perera, R.A.; Biskobing, D. Examining Bloom's Taxonomy in Multiple Choice Questions: Students' Approach to Questions. *Med. Sci. Educ.* **2021**, *31*, 1311–1317. [[CrossRef](#)] [[PubMed](#)]
- Kurdi, G.R. *Generation and Mining of Medical, Case-Based Multiple Choice Questions*; The University of Manchester: Manchester, UK, 2020.
- Zhao, W.; He, L.; Deng, W.; Zhu, J.; Su, A.; Zhang, Y. The effectiveness of the combined problem-based learning (PBL) and case-based learning (CBL) teaching method in the clinical practical teaching of thyroid disease. *BMC Med. Educ.* **2020**, *20*, 381. [[CrossRef](#)] [[PubMed](#)]
- Grainger, R.; Dai, W.; Osborne, E.; Kenwright, D. Medical students create multiple-choice questions for learning in pathology education: A pilot study. *BMC Med. Educ.* **2018**, *18*, 201. [[CrossRef](#)] [[PubMed](#)]
- Rakangor, S.; Ghodasara, D.Y.R. Literature Review of Automatic Question Generation Systems. *Int. J. Sci. Res. Publ.* **2015**, *5*, 1–5.

18. Zhang, R.; Guo, J.; Chen, L.; Fan, Y.; Cheng, X. A Review on Question Generation from Natural Language Text. *ACM Trans. Inf. Syst.* **2021**, *40*, 1–43. [[CrossRef](#)]
19. Rodriguez-Torrealba, R.; Garcia-Lopez, E.; Garcia-Cabot, A. End-to-End generation of Multiple-Choice questions using Text-to-Text transfer Transformer models. *Expert. Syst. Appl.* **2022**, *208*, 118258. [[CrossRef](#)]
20. Smith, P.E.; Mucklow, J.C. Writing clinical scenarios for clinical science questions. *Clin. Med.* **2016**, *16*, 142–145. [[CrossRef](#)]
21. Salam, A.; Yousuf, R.; Abu Bakar, S.M. Multiple Choice Questions in Medical Education: How to Construct High Quality Questions. *Int. J. Hum. Health Sci. IJHHS* **2020**, *4*, 79–88. [[CrossRef](#)]
22. Bowen, J.L. Educational Strategies to Promote Clinical Diagnostic Reasoning. *N. Engl. J. Med.* **2006**, *355*, 2217–2225. [[CrossRef](#)]
23. Hattie, J.; Timperley, H. The Power of Feedback. *Rev. Educ. Res.* **2007**, *77*, 81–112. [[CrossRef](#)]
24. Biggs, J. Aligning teaching for constructing learning. *High. Educ. Acad.* **2003**, *1*, 1–4.
25. Hadifar, A.; Bitew, S.K.; Deleu, J.; Develder, C.; Demeester, T. EduQG: A Multi-Format Multiple-Choice Dataset for the Educational Domain. *IEEE Access* **2023**, *11*, 20885–20896. [[CrossRef](#)]
26. Messick, S. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *ETS Res. Rep. Ser.* **1994**, *1994*, i-28. [[CrossRef](#)]
27. Prince, M. Epidemiology. In *Core Psychiatry*, 3rd ed.; Wright, P., Stern, J., Phelan, M., Eds.; Saunders Ltd. (Elsevier): Philadelphia, PA, USA, 2012; pp. 115–129. [[CrossRef](#)]
28. Messick, S. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am. Psychol.* **1995**, *50*, 741–749. [[CrossRef](#)]
29. Nunnally, J.C. Psychometric Theory—25 Years Ago and Now. *Educational Researcher. Educ. Res.* **1975**, *4*, 7–21. [[CrossRef](#)]
30. Kuder, G.F.; Richardson, M.W. The theory of the estimation of test reliability. *Psychometrika* **1937**, *2*, 151–160. [[CrossRef](#)]
31. Tarrant, M.; Knierim, A.; Hayes, S.K.; Ware, J. The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Educ. Today* **2006**, *26*, 662–671. [[CrossRef](#)] [[PubMed](#)]
32. Kim, K.H.; Zabelina, D. Cultural bias in assessment: Can creativity assessment help? *Int. J. Crit. Pedagog.* **2015**, *6*, 129–146.
33. Cizek, G.J. *Cheating on Tests: How to Do It, Detect It, and Prevent It*, 1st ed.; Routledge: New York, NY, USA, 1999. [[CrossRef](#)]
34. Masters, K. A Brief Guide to Understanding MOOCs. *Internet J. Med. Educ.* **2010**, *1*, 1–6. [[CrossRef](#)]
35. Larranaga, M.; Aldabe, I.; Arruarte, A.; Elorriaga, J.A.; Maritxalar, M. A Qualitative Case Study on the Validation of Automatically Generated Multiple-Choice Questions from Science Textbooks. *IEEE Trans. Learn. Technol.* **2022**, *15*, 338–349. [[CrossRef](#)]
36. Merchant, Z.; Goetz, E.T.; Cifuentes, L.; Keeney-Kennicutt, W.; Davis, T.J. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Comput. Educ.* **2014**, *70*, 29–40. [[CrossRef](#)]
37. Williams, J.B.; Wong, A. The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *Br. J. Educ. Technol.* **2009**, *40*, 227–236. [[CrossRef](#)]
38. Chen, B.; Bastedo, K.; Howard, W. Exploring Best Practices for Online STEM Courses: Active Learning, Interaction & Assessment Design. *OLJ* **2018**, *22*, 59–75. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.