Comparison of open and closed book test for admission in medical school [version 1]

Charline Cade¹, Jérémie Riou², Isabelle Richard³, Catherine Passirani⁴, Elisabeth Letertre⁴, Anne-Victoire Fayolle¹

¹General Practice Department
²MINT
³Faculté de Santé
⁴PluriPASS group

Abstract
This article was migrated. The article was marked as recommended.

Introduction: Students’ learning methods are highly influenced by the type of evaluation. Consequently, evaluation is a deciding step in the learning process. Multiple Choice Format Tests (MCQs) are almost exclusive in the admission to Health studies. PluriPASS suggests an educational reform of Health studies. This research aims at analyzing docimological features of Open-book Multiple Choice Format tests (OBT) compared to other usual tests used during the first year in Health studies curriculum (CC).

Methods: This educational research took place during PluriPASS year in Health studies curriculum at Angers University, during the academic period 2015-2016. The optional course « Disability and Health » (DH) was partly assessed by open-book tests with complex wording, requiring careful thought from students.

Results: Out of the 1161 students enrolled in first year curriculum, 190 have chosen the DH. For the students who followed the DH, the CC and OBT distributions are respectively a Skewness score at -0.11 and -0.12 and a Kurtosis score at -0.9 and -0.22. Bland Altman test or Deming method demonstrate a concordance between both method.

Conclusion: Distribution characteristics of OBT are satisfactory and allow to consider introducing this method with the objective of promoting reflection and depth learning.
Keywords
Study skills, Medical education research, Evaluation, Medicine

Corresponding author: Charline Cade (charline.cade@gmail.com)
Competing interests: No competing interests were disclosed.
Grant information: The author(s) declared that no grants were involved in supporting this work.
Copyright: © 2018 Cade C et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
How to cite this article: Cade C, Riou J, Richard I et al. Comparison of open and closed book test for admission in medical school [version 1] MedEdPublish 2018, 7:25 https://doi.org/10.15694/mep.2018.0000025.1
First published: 25 Jan 2018, 7:25 https://doi.org/10.15694/mep.2018.0000025.1
Introduction

Admission procedures have been considered by many authors as one of the most important step in the design of a medical curriculum (Patterson & Al., 2016). Indeed in most higher education systems, the proportion of students that will complete medical studies after admission is very high and thus the admission procedure is crucial in shaping the working force for the health system. A large variety of procedures have been described and are currently used by medical schools in the world. (Adam, Dowell, & Greatrix, 2011; Fayolle & Al., 2016; Groves, Gordon, & Ryan, 2007)

The admission process can take place at different points of the curriculum and it is roughly possible to distinguish undergraduate entry in medical schools, recruiting students immediately after secondary school and graduate entry curricula, recruiting students who have previously completed a bachelor course or equivalent. Some institutions offer both opportunities. (James, Fergusson, Powis, Symonds, & Yates, 2008; Elliott & Epstein, 2005)

French regulations allow all students having completed secondary education to apply to university in most fields, including health sciences (Arrêté du 28 octobre 2009). Students are admitted in a multidisciplinary first year program designed for large numbers of students and compulsory to apply as a second year student for medicine, pharmacy, dentistry and midwifery studies. Therefore the real bottleneck for admission to health studies is the end of first year competitive exam, which can be compared to the admission procedure of other higher education systems.

The processes used for screening, and finally selecting medical students vary from one institution to another. Most systems include the assessment of academic achievements, either by some measurement of secondary education achievements (such as grade point averages, marks for final secondary education exams, etc...), or by specific entrance exams assessing previously acquired knowledge. Some systems include tests specifically designed for recruiting medical students, sometimes on a nationwide basis, such as the MCAT (Medical college admission test) (Pau & Al., 2013) or the UKCAT (United Kingdom Clinical aptitude test) (Reiter & Eva, 2005). In many medical schools the application procedure includes a written essay of the motivations, support letters, and attention to prior activities in the field of health or community support activities. Some medical schools include an interview, usually for a subset of selected students. Several interview patterns have been described among which the Multiple Mini Interviews developed by Macmaster University (Eva, Rosenfeld, Reiter, & Norman, 2004) and widely applied in Canadian medical schools (Pau & Al., 2013).

Designing the admission procedure involves several steps and several stakeholders.

The first step determines the applicant’s characteristics that are considered as relevant for admission. The trend is toward a shift from assessment of purely academic characteristics to a more competency-based assessment, including for instance communication skills or ethical reasoning (Reiter & Eva, 2005). The second step is designing tests or procedures that accurately assess these characteristics. Since the number of applicants is usually much higher than the number of available positions, the ability of the procedure to discriminate between students appears more determinant than the ability to set a minimal level. In other words many students applying for medical school will perform fairly well on many procedures and the challenge is to distinguish very good candidates from good candidates.

The stakeholders appear to have different priorities in designing the admission procedures9. The health system, and for instance the representatives of the future patients have expressed their concern that the admission procedure should ensure that the students have the relevant communication and ethic reasoning skills. The students are extremely sensitive to the fairness of the procedure; and the universities need to control the overall organization costs (Hissbach, Sehner, Harendza, & Hampe, 2014).

In France, the first year competitive exam consists exclusively in written exams, including multiple choice questions for a set of basic science subjects and an essay for social sciences. In all French universities these exams are closed book exams. The multiple choice questions have usually fairly short stems and assess mainly basic recall. These procedures have been accused of selecting “memory freaks”, and being irrelevant when it comes to assessing the ability to deal with uncertainty (Edutopia, 2017). Despite a number of criticisms these procedures have been maintained because of their easy and cheap implementation, and because they are well accepted by students. Students associations consider them to be fair even if they might admit that they are incompletely relevant.

The French ministry of higher education has issued in 2013 a call for proposal and some universities have been selected to implement different admission procedures. The admission program (called PluriPASS) implemented at the university of Angers has been completely redesigned and consists in a core program and elective self selected modules. These modules represent only a small proportion of the final mark and have been used to develop and test different learning strategies and
different assessment procedures. One of these modules (Disability and Health (DH) module) consists of a 20 hours online content presenting concepts and issues related to disability. The assessment of this module is an open book exam of 15 multiple choice questions with long stems, completed in one hour.

Open book tests (OBT) have been widely described and studied (Feller, 1994). Open book exams could favor deep learning strategies and could be a method to assess how the student uses available knowledge rather than memorizing large amounts of data. This could be particularly relevant in a world of growing available information. Therefore open book multiple choice questions could be a strategy that combines low cost, reproducibility and “fairness” of multiple choice questions as well as the ability to investigate deeper cognitive strategies. OBT have also showed a trend towards better results for all or most students (Feller, 1994; Tarrant, Ware, & Mohammed, 2009; Stalnaker & Stalnaker, 1934; Heijne-Penninga, Kuks, Hofman, & Cohen-Schotanus, 2010). This could be a limit in using OBT for admission procedures if the distribution of the marks is such that distinguishing excellent from good students on this basis becomes difficult.

The objective of this study is to analyze the results on an open book test in a first year student population, and compare them to closed book test (CBT) results.

Methods
Population and Method
The study concerns the population of students enrolled in the first year program during the 2015-2016 academic year at the University of Angers (France) who selected the “Disability and Health” module (DH Module).

The first year program included a core curriculum of 12 subjects. Each subject was assessed by a CBT multiple choice question (MCQ) exam and for two subjects in the field of social sciences by MCQs and a short essay. Students obtained a mark which was the sum of the 12 marks (maximum 1200 points).

Each student also selected three elective modules. These modules were assessed by a range of methods. The DH module assessment included a one hour open book exam. The results obtained for the elective modules were considered on a fail/pass basis and a “pass” result provided an additional 20 bonus points.

The total mark (including the sum of the core curriculum marks and the possible additional elective bonuses) was used to rank the students for admission to the second year of medical studies. Based on this result the 63 students with the highest marks were admitted. The 119 following students were invited for additional interviews. 64 of these students were admitted after the interviews.

Age, gender, core curriculum mark and final admission status were obtained for all students.

For the sub-population of “DH module students” the marks for the OBT was extracted.

Data processing
The characteristics of the sub-population of DH students were compared to the overall population by student test, or test and Fisher’s exact test, as appropriate.

The core-curriculum marks were expressed as a fraction of 20 (sum of the 12 modules (max 1200) /60), which is the usual expression of marks in the French academic system.

The distribution of core-curriculum marks, and DH module book test were established. Each distribution was described by the Skewness and Kurtosis coefficients (Groeneveld & Meeden, 1984). The Skewness coefficient describes how symmetrical the distribution is and the Kurtosis coefficient how sharp (or flat). The value of both coefficients is 0 for a normal distribution. This analysis was performed to obtain data as to the ability of each method to discriminate and rank students. For instance, a very high Kurtosis coefficient would mean that the distribution is very sharp and that many students obtain the same mark, which is inappropriate if the aim is to rank students.

The correlations between the marks obtained by a given student for the core curriculum and for OBT of the DH module were analyzed by the graphical method of Bland and Altman (Martin & Altman, 1986). This method is used to determine whether one measure is systematically biased against another and whether both methods are consistent, by determining
the percentage of differences within an interval of two standard deviations. Finally, considering that these marks are used to decide admission, we defined the number of students who ranked in the top 20% for each method.

**Results**

**Population**

1312 students entered in September 2015 the first curriculum year. 151 dropped out during the first month and did not show up for the first exam. (Figure 1)

DH: “Disability and Health” optional course

201 students selected the DH elective module; compared to the overall population these students are more often women (88% versus 70.5%, p <0.05) and do not differ on other characteristics such as their achievements for the core-curriculum modules or their final admission status. (Table 1)

![Students flow chart](image)

*Figure 1. Students flow chart.*
Thirty-four students having selected the DH module dropped out and did not show up for the DH exam. These students also achieved very poor results on the core curriculum modules as they did not show up for some or all exams and were removed from further analysis.

### Marks distribution

The distribution of the marks was analyzed for the core-curriculum mark and for OBT and is presented in the (Figure 2). All marks are expressed as a fraction of 20.

#### Table 1. Students’ characteristics in PluriPASS. Comparison of the sub-population of DH students to the overall population.

|                        | DH module | Population | p-value* |
|------------------------|-----------|------------|----------|
|                        | n=190     | n=1161     |          |
| Male gender            | 23(12)    | 342(29)    | 0.005    |
| Core Curriculum mean mark (m ± sd) | 10.24 ± 4.32 | 10.97 ± 4.49 | 0.206 |
| Admission to the second year of medical studies | 35(18) | 261(22) | 0.210 |

*Data presented as n(%), where applicable.

*Based on t-test or Fisher’s exact test as appropriate.

*p-value significant (i.e., <0.05) indicated in bold.

#### Figure 2. Superposition of distribution curves: core-curriculum (blue) and Open-book Test (pink).

Thirty-four students having selected the DH module dropped out and did not show up for the DH exam. These students also achieved very poor results on the core curriculum modules as they did not show up for some or all exams and were removed from further analysis.
The mean (10.88 in core curriculum vs. 12.77 in OBT), median (10.55 vs. 12.82), variance (11.89 vs. 7.02), skewness (-0.11 vs. -0.12) and kurtosis coefficients (-0.9 vs. -0.22) are presented in Table 2.

The mean mark is significantly higher for the OBT than the average core-curriculum mark (12.77 vs. 10.88, p<0.05).

Homogeneity of methods
The Bland Altman analysis was applied to compare the homogeneity of the OBT and core curriculum marks (CC) (Figure 3). As previously stated the average observed bias is 1.89. The graph shows that few points stand outside the amenity limits set at two standard deviations around the observed systematic difference.

Finally we determined which students belonged to the top 20% for both methods and determined the conditional probability of belonging to the top 20% on the OBT knowing that a student belongs to the top 20% of the core curriculum and vice-versa. This shows that among the top 20% for the CC mark, 50% also belong to the top 20% for the OBT (and 50% do not).

Discussion
Our results show that students obtain higher marks on OBT of the elective module than on the CBT of the core curriculum. The distribution of OBT marks remains fairly symmetrical as attested by the skewness coefficient close to zero. The distribution is sharper than for the CBT but remains flatter than a normal distribution. 10% of the students belong to the

| Table 2. Marks characteristics of core curriculum and Open book test |
|---------------------------------------------------------------|
|                  | Mean | Variance | Skewness | Kurtosis |
| Core curriculum  | 10.88| 11.39    | 0.11     | -0.9     |
| Open book test   | 12.77| 7.02     | -0.12    | -0.22    |
The results of the students were better on OBT than on the core curriculum, despite the fact that the “bonus” system did not encourage students to get marks above the pass threshold. This has been previously described for OBT (Chan & Mui, 2004; Skidmore & Aagaard, 2004). Authors have shown that this is not due to more efficient learning method when preparing for OBT (Dickson & Bauer, 2008), and that OBT is not relevant if the objective is to assess the correct memorization of core knowledge, can lead students to reduce their workload and can decrease long-term retention.

The absolute measure is important if the objective of the assessment is to set the minimal level that students should achieve (for instance to get a “pass” result). In such contexts the pass threshold could be set at a higher level for OBT than for CBT.
The distribution of the marks for the OBT was not very different from the distribution for other tests, and the Kurtosis coefficient, which assesses how flat the curve is was lower than for a normal distribution. This parameter is crucial if the objective is to rank students and therefore spread the distribution. The intuitive concern that OBT could lead to all students achieving high marks making it impossible to rank them effectively was not confirmed and we believe that this preliminary result makes it safe to introduce some proportion of open book testing in the core curriculum.

In admission procedures, one of the key issues is whether the student is within the top 20% who are offered admission. Comparing the core curriculum marks and OBT, 10% of the students fall in the top 20% in both cases, and 10% for one method but not for the other. This could indicate that the abilities assessed by both methods are indeed different.

Open book multiple choice questions are more frequently rich context questions and this was the case in our study. Indeed the teacher is fully aware that the students will bring their notes and anticipates that testing purely factual knowledge is not relevant, and therefore tends to write complex questions requiring the students reasoning. This can also be done for close book tests if the teachers are clearly instructed to do so. One interpretation of our results could be that the open book rich context questions did indeed assess reasoning, more or differently from the questions of the core curriculum, but it is not possible to determine if this is due to the open book testing or to the rich context. Open book test appears in fact as an institutional backdoor to motivate teacher to write rich context questions. Further work could differentiate rich context open book tests and rich context closed book tests.

Notes On Contributors

Authors’ names: Charline Cade¹; Jérémie Riou²; Isabelle Richard³; Catherine Passirani⁴; Elisabeth Letertre⁴; Anne-Victoire Fayolle¹

Affiliations: (department, university, city, country)

¹General Practice Department, University of Angers, France
²MINT, UNIV Angers, INSERM 1066, CNRS 6021, Bretagne Loire University, IBS-CHU, Angers, France
³Faculté de Santé, University of Angers, France
⁴PluriPASS group, Faculté de Santé, University of Angers, France

Authors’ contributions: All authors read and approved the final manuscript

Declarations

The author has declared that there are no conflicts of interest.

Appendices

List of abbreviations
CBT: Closed book test
CC: Health studies curriculum
DH: “Disability and Health” optional course
MCAT: Medical college admission test
MCQ: Multiple choice format test
OBT: Open-book Multiple choice format test
UKCAT: United Kingdom clinical aptitude test
USMLE: United States Medical Licensing Examination

Declarations

Ethics approval and consent to participate: Not applicable
Consent for publication: Not applicable
Availability of data and material: Not applicable
Competing interests: None
Funding: None
Bibliography/References

Adam, J, Dowell, J, Greatrix, R (2011). Use of UKCAT scores in student selection by U.K. medical schools, 2006-2010. *BMC Med Educ.* 11: 98.

**Reference Source**

Arrêté du 28 octobre 2009 relatif à la première année commune aux études de santé.

Bloom, BS, Others, And (1971). Handbook on Formative and Summative Evaluation of Student Learning. McGraw-Hill Book Company, 330 West 42nd Street, New York, New York 10036 (11).

Brightwell, DR, Daniel, MI-H, Stewart, DA (2004). Evaluation: is an open book examination easier? *BioSci Educ.* May 1; 3(1): 1–10.

Chan, MY, Mui, KW (2004). The use of open-book examinations to motivate students: a case study from Hong Kong. 2004 [cited 2017 Jul 18]; Available from:

**Reference Source**

Dickson, KL, Bauer, JJ (2008). Do Students Learn Course Material during Crib Sheet Construction? *Teach Psychol.* Apr 1; 35(2): 117–20.

Elliott, SL, Epstein, J (2005). Selecting the future doctors: the role of graduate medical programmes. *Intern Med J.* Mar; 35(3): 174–7.

Feller, M (1994). Open-book testing and education for the future. *J High Educ.* Jan; 65(4): 389–96.

Fayolle, A-V, Passirani, C, Letertre, E, Ramond, A, Perrotin, D, Saint-André, J-P, et al. (2016). (Predictive validity of selection process in medical school, a systematic review of the literature). *Presse Med.* Paris Fr. 1983. May: 45(5): 483–94.

Feller, M (1994). Open-book testing and education for the future. *Stud Educ Eval.* Jan 1; 20(2): 235–8.

Green, M, Argoft, N, Encandela, J (2016). Test anxiety and United States Medical Licensing Examination scores. *Clin Teach.* Apr; 13(2): 142–6.

Groeneveld, RA, Meeder, G (1984). Measuring Skewness and Kurtosis. *J R Stat Soc Ser Stat.* 33(4): 391–9.

Heijne-Penninga, M, Kuku, JBM, Hofman, WHA, Cohen-Schotanus, J (2008). Influence of open- and closed-book tests on medical students’ learning approaches. *Med Educ.* Oct; 42(10): 967–74.

Heijne-Penninga, M, Kuku, JBM, Hofman, WHA, Cohen-Schotanus, J (2010). Influences of deep learning, need for cognition and preparation time on open- and closed-book test performance. *Med Educ.* Sep; 44(9): 884–91.

Hisbach, JC, Sehner, S, Harendza, S, Hampe, W (2014). Cutting costs of multiple mini-interviews - changes in reliability and efficiency of the Hamburg medical school admission test between two applications. *BMC Med Educ.* Mar 19; 14: 54.

Iliannidou, M (1997). Testing and life-long learning: Open-book and closed-book examination in a university course. *Stud Educ Eval.* Jan 1; 23(2): 131–9.

James, D, Ferguson, E, Powis, D, Symonds, L, Yates, J (2008). Graduate entry to medicine: widening academic and socio-demographic access. *Med Educ.* Mar; 42(3): 294–300.

Martin Bland, J, Altman, D (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet.* Feb 8; 327(8476): 307–10.

Moores, R, Jensen, PA (2007). Do Open-Book Exams Impede Long-Term Learning in Introductory Biology Courses? *J Coll Sci Teach.* Jul; 36(7): 46–9.

Patterson, F, Knight, A, Dowell, J, Nicholson, S, Cousans, F, Cleland, J (2016). How effective are selection methods in medical education? A systematic review. *Med Educ.* Jan; 50(1): 36–60.

Pau, A, Jeevaratnam, K, Chen, YS, Fall, AA, Khoo, C, Natarajah, VD (2013). The Multiple Mini-Interview (MMI) for student selection in health professions training - A systematic review. *Med Teach.* Dec; 35(12): 1027–41.

Reiter, HI, Eva, KW (2000). Reflecting the relative values of community, faculty, and students in the admissions tools of medical school. *Teach Learn Med.* 17(1): 4–8.

**Reference Source**

Skidmore, RL, Aagaard, L (2004). The relationship between testing condition and student test scores. *J Instr Psychol.* Dec; 1; 31(4): 304–14.

Stalnaker, JM, Stalnaker, RC (1934). Open-Book Examinations. *J High Educ.* Mar 1; 5(3): 117–20.

Tarrant, M, Ware, J, Mohammed, AM (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC Med Educ.* Jul 7; 9: 40.

Whitley, J, Bernard, E (1996). Does cheating help? The effect of using authorized crib notes during examinations. *Coll Stud J.* 30(4): 489–93.
Open Peer Review

Migrated Content

Version 1

Reviewer Report 01 May 2019

https://doi.org/10.21956/mep.19495.r28808

© 2019 Masters K. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ken Masters
Sultan Qaboos University

This review has been migrated. The reviewer awarded 3 stars out of 5

A reasonably interesting paper dealing with a comparison of open-book testing as opposed to closed-book testing for entrance tests to medical schools. The authors have properly identified the central problem and the contextual literature background, have performed a detailed analysis of the data (with a strong sample size), and then have related their Results back to the literature in the Discussion. The paper has many strengths, but, unfortunately, is undermined by some careless work in the presentation. Some things to consider: • It is a pity that the authors elected to not have a separate Conclusion, and did not use the “Take Home Messages” option of the journal. In this paper, a large amount of information has been given, and the Discussion is very detailed. This is a strength of the paper. An associated weakness, however, is that many readers can become overwhelmed, and not see the overall impact of the study. A succinct Conclusion and bullet-point “Take Home Messages” helps significantly to clarify the central messages of the paper. • Quite a few small errors, such as errors of language, upper-case and lower-case inconsistency, lack of commas where they should be, typographical errors, incorrect word order, etc. These really should have been fixed before submission, as they serve only to irritate the reader. • The conclusion of the Abstract is really difficult to follow, and some greater care should have been taken on making sure that this is very clear. So, a useful study, but the authors really should have taken greater care in the presentation of their material.

Competing Interests: No conflicts of interest were disclosed.

Reviewer Report 25 January 2018

https://doi.org/10.21956/mep.19495.r28807
This review has been migrated. The reviewer awarded 4 stars out of 5

I enjoyed reading this article about an open book exam as part of the admission process in a French University. The authors also provide a good overview of student admission to medicine. I also obtained an overview of the admission process to the French system of medical education. Open book tests can encourage critical thinking and application of information by the students. The authors describe how in the modern era technology has made it easier to access information. Hence, learning to interpret and apply information may be a very important skill. Modern doctors in practice have ready access to electronic resources which can quickly retrieve information. The authors can provide ‘Take Home Messages’ toward the end of their manuscript. A brief description about the University of Angers and the medical school will be helpful to international readers. The standard of written English in the manuscript is good, though corrections/modifications can be carried out in a few places. Admission tests using knowledge-based multiple choice questions are widely used in many institutions and countries mainly due to ease of standardization, perceived objectivity and the need for lesser resources. Many institutions are increasingly using other methods to evaluate applicants. This paper will be of interest to medical educators and especially to those with a special interest in student admissions.

**Competing Interests:** No conflicts of interest were disclosed.