

WEB PAPER

Is a web-based concordance test feasible to assess therapeutic decision-making skills in a French context?

LOUIS SIBERT^{1,2}, ROCH GIORGI³, BADISSE DAHAMNA¹, JEAN DOUCET², BERNARD CHARLIN⁴ & STEFAN J. DARMONI¹

¹Rouen University, France, ²Rouen University Hospital, France, ³Aix-Marseille University, France, ⁴University of Montreal, Canada

Abstract

Background: At the end of undergraduate medical curriculum, a written simulation-based examination is used in France to assess therapeutic decision-making skills and to rank students for the purpose of matching their training specialties. However, this examination based on a single assessment method remains a subject of debate.

Aim: To study the feasibility of a web-based Concordance test for therapeutic decision-making assessment.

Methods: A 12 clinical-case Concordance test was developed based on objectives for the undergraduate training program. The test was administered on line to candidates with different levels of clinical experience. Fifteen therapeutic teachers constituted the reference panel. Data analysis included analysis of variance, *post-hoc* test, and Cronbach's alpha.

Results: One hundred and seventy participants (113 students, 34 residents, 23 physicians) fully completed the free-access test on line with no technical problems. Differences between the mean scores for groups were significant ($p < 0.001$). Significant differences occurred between fourth year students and residents ($p < 0.001$), fourth year students and physicians ($p = 0.001$). No difference was found between residents and physicians. Reliability coefficient was 0.67.

Conclusion: A web-based Concordance test in the field of therapeutic decision-making was considered feasible in a French learning environment. Further research is warranted to determine its usefulness as a part of the National Examination.

Introduction

The primary goal of medical teaching programs is the acquisition of clinical competence. Therapeutic care represents a major outcome of clinical competence. Although a sound theoretical knowledge based on Evidence-Based-Medicine (EBM), as well as clinical and interpersonal skills are vital for a physician; decision-making skills are a crucial component of therapeutic care competence. Traditional tools for assessing clinical competence can accurately and reliably test theoretical knowledge from EBM (i.e. rich-context, multiple choice questions) and clinical as well as interpersonal skills (i.e. Objective Structured Clinical Examination) (Van der Vleuten 1996). However, in daily practice, treatment decision-making requires both knowledge and experience and is a patient-oriented decision-making process. In each clinical encounter, the patient's capacity to accept treatment, understanding and follow-up must be carefully considered in order to achieve effective patient therapeutic care outcome. Therapeutic reasoning in the medical profession is much more than a simple application of knowledge, rules, and principles. Treatment decisions must be individualized for each patient, at each stage of their disease. When confronting a patient, fuzzy, frequently missing or contradictory data must be retrieved in order to solve the therapeutic problem.

Practice points

- Therapeutic decision-making is a major learning and assessment objective for the second part of the undergraduate curriculum in France.
- The selection process for students is based on a single assessment tool (patient management problem) and is controversial.
- A Web-based Concordance test is feasible in the French context with accurate reliability.
- Further research is required before including Concordance test as a part of the French National undergraduate Examination.

Therefore, a significant part of therapeutic decision-making skills stems from the ability to solve ill-defined problems where EBM cannot be applied.

The Concordance test, which is based on cognitive psychology script theory (Charlin et al. 2000), may provide a useful method to build a theory-based assessment tool containing tasks as complex and ambiguous as that found in clinical practice. The Concordance test was developed with the purpose of probing whether knowledge of examinees is

Correspondence: Louis Sibert, Department of Urology, Rouen University Hospital, 1, rue de Germont, 76031 Rouen, Cedex, France. Tel: (33) 2 32 88 81 73; fax: (33) 2 32 88 82 05; email: louis.sibert@chu-rouen.fr

effectively organized for clinical decision-making. It is designed to be added to existing tools, not substituted by them. The test has 3 basic principles (Charlin & Van der Vleuten 2004): (1) examinees are faced with a challenging authentic clinical situation in which several options are relevant, (2) responses follow a Likert scale that reflects script clinical reasoning theory, and (3) scoring is based on the aggregate scoring method to take into account the variability of the clinical reasoning process among experts. Previous studies using Concordance tests with different contents have demonstrated reliability and validity for this tool (Brailovsky et al. 2001; Charlin & Van der Vleuten 2004; Charlin et al. 2007; Gagnon et al. 2006; Meterissian et al. 2007).

Therapeutics training programs were developed a few years ago in French Medical schools with the aim of teaching both general curative care (i.e. the rational use of drugs as well as non-pharmacological treatments) and clinical decision-making. These training programs were implemented with the perspective of global and multidisciplinary patient approach. In the French medical educational system, learning objectives for the second part of the undergraduate curriculum (years 3–6) are dedicated to teaching decision-making skills with the goal of acquisition of curative care competence and to adapt it to the clinical context (Segouin et al. 2007). At the end of the undergraduate curriculum, a National Examination is used to rank students for the purpose of matching to specialty training. This National Examination is based on a common program of 345 questions (i.e. question 302: Acute diarrhea of adults and children). Therefore, the individual teaching programs presented in each French medical school are very similar. This national ranking assessment is written patient management problem-based. The introduction of a method of selection, which is based on a single assessment tool, is a French peculiarity and currently the subject of several debates. First, it is well established in the literature that test formats based on written simulations of clinical problem solving, such as patient management problem, have repeatedly shown that experienced clinicians often perform slightly better and sometimes worse than clinicians with lower levels of experience (Van der Vleuten 1996); Second, for clinical reasoning assessment, it is difficult to reach a consensus on a unique 'good answer' for each item among jury members (Charlin et al. 2002). Furthermore, numerous French medical teachers argue that this examination strongly promotes rote learning strategies among students rather than deep and reflective learning. Moreover, it has been shown that, in order to maximize their chances of being highly ranked, students frequently attend supplementary training education outside medical school at their own expense (Segouin et al. 2007). Based on these data, French medical teachers are not comfortable with this assessment tool and are currently trying to improve its usefulness (Segouin et al. 2007).

To date, Concordance test used to specifically evaluate treatment decision-making skills has not been previously reported. In this context, the purpose of this study was to measure the feasibility of concordance test when applied in the domain of therapeutic decision-making throughout the French environment. In order to answer the research question, participants were recruited from the entire country and were

identified according to their level of medical training (students, residents, board certified practitioners). In order to facilitate administration of the Concordance test on a very large number of candidates over a short period of time, we promoted an on line assessment course of therapeutic decision skills using the Concordance test. This study was conducted in the field of emergency care and internal medicine under the French Medical Virtual University (FMVU) consortium umbrella (Le Beux et al. 2000; FMVU 2008). The FMVU consortium of 31 French medical schools out of 32 has the objective of sharing experiences throughout the entire country using information and communication technologies (ICT) to support new teaching perspectives for medical students (Sibert et al. 2006a). FMVU is not a new medical school and has not added any fees for medical students.

Methods

Construction of the test

A Concordance test for Therapeutics has been developed according to the methodology previously described (Charlin & Van der Vleuten 2004), based on major educational objectives of therapeutics French undergraduate training programs. French therapeutics training programs are based on a multidisciplinary approach. Concordance test content was built based on this approach. In order to avoid unnecessary focus on one single specialty, we arbitrary decided to target our attention on emergency care as well as internal medicine. When preparing clinical vignettes, an attempt was made to keep clinical scenarios authentic as well as to require reasoning skills and some clinical experience. The description of each situation had to be complex enough to be challenging for the level of training that we wanted to assess (i.e. undergraduate level). Each vignette was followed by a series of related items focused on treatment strategies. The material obtained at this preliminary stage was then assessed by two teachers of therapeutics in order to verify whether the test in fact addressed realistic therapeutic care dilemmas and if it tested decision-making capacities. This process allowed us to retain the optimal questions for the Concordance test administered to participants. A total of 12 clinical cases, each of them followed by a series of three related questions were included in the database and constituted the final Concordance test on line. Examples of clinical vignettes with items are illustrated in Table 1.

Development of the web site

The development process of the Web site using Concordance test has been described in a previous report. We have recently introduced a Web-based Concordance test in the field of urology. To our knowledge, this study was the first reported Web-based assessment course with the concordance test (Sibert et al. 2006b). This process adheres to the main principles of practical guidelines for developing an effective educational website (Cook & Dupras 2004) and is based on the traditional three-tier architecture used in Web database applications.

Table 1. Examples of clinical vignettes and format of items used for the concordance test.

Clinical vignette no. 1: You are called to the home of a 36-year-old plumber who presents with the first episode of acute lumbago on the right side, 24 h previously, following a period of heavy lifting.		
If you were considering to prescribe	And then you discover	The relevance of this treatment becomes
Bi-Profenid® 150 mg per os 2/days	Previous medical history of chronic sinusitis	-2 -1 0 +1 +2
Profenid® 100 mg, 1 injection IM	He had a urinary tract infection the previous week treated with Bactrim	-2 -1 0 +1 +2
Rest in bed for 13 days with Efferalgan Codeine, 2 per day	Temperature of 38°2	-2 -1 0 +1 +2
Clinical vignette no. 2: A 79 year old patient was hospitalized in the emergency department for pain resistant to treatment associated with a major alteration in his general status within a context of a developing 30-year history of sclerodermia.		
If you were considering to prescribe	And then you find	The relevance of this treatment becomes
Morphine treatment	A very great reluctance of the patient to this type of treatment	-2 -1 0 +1 +2
No therapeutic treatment with a period of watchful waiting	Total refusal by the patient of any treatment whatsoever	-2 -1 0 +1 +2
Antidepressant treatment	There are no signs of depression	-2 -1 0 +1 +2
Clinical vignette no. 3: A 58-year-old man has high blood pressure (160/90 mmHg) treated to date with a Beta-blocker. This treatment should be discontinued due to an atrioventricular block dysfunction. An alternative antidepressant treatment is required.		
If you were considering to prescribe	And then you find	The relevance of this treatment becomes
An enzyme conversion inhibitor	A creatinine clearance of 45 mL/mn	-2 -1 0 +1 +2
An central antihypertensive drug	A benign prostatic hyperplasia treated by an alpha-blocker	-2 -1 0 +1 +2
A calcium antagonist	Non-insulin-dependent diabetes mellitus	-2 -1 0 +1 +2

Notes: -2: contra-indicated; -1 not useful or even detrimental; 0: nor less no more useful; +1: useful; +2: necessary or absolutely necessary.

Scoring process

The scoring process is derived from the aggregate scoring method (Charlin et al. 2002). This system takes into account the range of potential answers and allows for the variability in reasoning process that experienced clinicians show when confronted with ill-defined, complex situations. Credits for each item are derived from the answers given by a panel of reference. For each item, candidates' answers received a credit mark corresponding to the proportion of panel members who selected it. The maximum score for each item was 1 for the modal answer. Other panel members' choices received a partial credit. Answers not chosen by panel members received 0. To obtain this proportional transformation, the number of panel members who had provided an answer on the Likert scale was divided by the modal value for the item. For example, if on an item, six panel members (out of 10) had chosen response +1, this choice received a score of 1 point (6/6). If three panel members had chosen response +2, this choice received a score of 0.5 (3/6), and if one panel member had chosen response 0, this choice received a score of 0.16 point (1/6). The total score for the test was the sum of credits obtained on each item, which in the end was transformed to obtain a maximum of 100.

Reference panel

One of the features of French therapeutics training programs is that teachers are from various medical disciplines. Taking into account this multidisciplinary approach of the patient, inclusion criteria in order to be a reference panel member had to be (1) faculty member and (2) internist or emergency care practitioner. Previous studies have demonstrated that recruiting more than 10 members in a reference panel presents a reliable assessment of clinical reasoning with the Concordance test whereas using more than 20 members shows only a

marginal benefit in terms of psychometrics properties (Gagnon et al. 2005; Charlin et al. 2007). Taking into account these data, 15 internists and emergency care practitioners, all members of the Association Pédagogique Nationale pour l'Enseignement de la Thérapeutique (APNET) and all involved in therapeutics training programs in their faculties, constituted the reference panel. They were asked to fill out the test online individually, under the same conditions as future candidates.

Candidates

In order to assess psychometric properties of the Concordance test, different groups of participants were identified according to their level of medical experience: students, residents, and board certified clinicians. Participants were recruited on a voluntary basis by e-mail from the APNET and the FMVU Web sites. Inclusion criteria to participate in the Concordance test were: for the Board certified clinicians, to be member of the APNET, for the residents, to be involved in an internal medicine or emergency care training program from one of the FMVU medical schools and for the students, to be involved in one of the FMVU medical schools. As therapeutics-teaching programs are implemented at the beginning of the fifth year, students were also identified according to their level of medical training (fourth, fifth or sixth year). Each participant, as well as each reference panel member, was assigned by a unique identification number in order to enter the Web site.

Statistical analysis

Item scores and total scores for each participant were computed and statistical analyses were performed using the SPSS software version 13.0 (SPSS Institute Inc., Chicago, IL). Descriptive statistics of the participants' scores on the Concordance test were performed (group's mean scores with

http://www.chu-rouen.fr/tcsecn/

Université Médicale Virtuelle Francophone

T.C.S. E.C.N. (Test de Concordance de Script pour l'Examen Classant National)

Accueil Aide Inscription Test Resultats Statistiques Liens

Ce questionnaire vise à explorer les connaissances utilisées pour résoudre les problèmes en pratique clinique.

Il est bien évident :

- Que votre participation à cette étude restera anonyme.
- Que l'objet de cette étude est d'évaluer la pertinence du questionnaire et non la compétence des participants.
- Que les résultats de cette étude ne seront utilisés que pour la recherche statistique et n'entreront pas en compte pour une quelconque notation des participants.

Le questionnaire vous place dans des contextes cliniques spécifiques décrits dans des vignettes. Pour chaque question, une information spécifique vous est fournie et différentes options de diagnostic, d'investigation ou de traitement vous sont proposées. Il vous est demandé quel effet l'option présentée a pour vous selon l'information disponible (voir l'Article). Les informations, ainsi que les options sont introduites au hasard. Chaque question est indépendante. Il ne s'agit pas de déterminer l'effet du recueil d'un ensemble d'éléments d'information, mais plutôt de déterminer l'effet d'une option isolée (selon un élément d'information spécifique).

Afin de pouvoir débuter le test, il vous faut au préalable un code participant. Vous pouvez l'obtenir via l'onglet **Inscription**.

Contacts : Louis Sibert (membre du bureau pédagogique de la faculté) - Stéfan Darmoni (professeur d'informatique médicale) - Badisse Dahanna (ingénieur informaticien) - Jacques Weber (directeur de la commission pédagogie du GIP UMVF) - Mis à jour le 11/03/2004. Ce site adhère aux principes de qualité du NetScoring - Copyright © 2004 CISMéF -

Main modules of the concordance test on line:

- registration
- test
- individual score
- global scores

Figure 1. Illustration of the concordance test on line home page (URL: <http://www.chu-rouen.fr/tcsecn>).

the associated 95% confidence interval), followed by a 1-way analysis of variance to evaluate differences between the group's mean scores. To evaluate the presence of a significant statistical difference, a $p < 0.05$ was considered as significant. A Bonferroni correction procedure was then used to accurately determine which score differences were significant between groups of participants. To evaluate the presence of a significant statistical difference, an adjusted $p < 0.05/6 = 0.0083$ was considered significant. Reliability of the examination was assessed via the Cronbach's alpha coefficient. The Concordance test is a format of case-based assessment, each question following the vignettes is dependent on the content of each clinical vignette. Therefore, in this study, the clinical vignettes were used as units of reliability analysis and reliability analyses were performed using the 12 clinical vignettes.

Results

The Concordance test was placed on line at the Rouen University Hospital Web site (URL: <http://www.chu-rouen.fr/tcsecn>) under the umbrella of the FMVU consortium. The evaluation system can be assessed via any computer system with a standard Web browser. Contrary to the Concordance test in urology, which was limited to French urologists (Sibert et al. 2006b), the concordance test in therapeutics was open to any members of the APNET and any students or trainees from medical schools members of the FMVU consortium. The home page of the test (Figure 1) contains one module to register as a candidate or as member of the reference panel, another to pass the test, and another to obtain the global scores. The

home page contains a summary of the concordance test principles as well the instructions for the participants.

The Concordance test has been operational since October 2005. The inclusions for this study were stopped on June 2007. During this 20-month period, a total of 15 reference panel members and 251 candidates had passed the Concordance test on line. As demonstrated in our previous study (Sibert et al. 2006), most of the participants primarily came from 'Ile de France' area ($N = 125$) and from 'Normandy' geographical area ($N = 75$); the other participants ($N = 51$) came from various French areas. Mean examination time was approximately 50 minutes. During this period, no technical problems were encountered on the web site. Nevertheless, 81 candidates (32%) did not answer all the items. A total of 16 fourth-year students, 48 fifth-year students, 49 sixth-year students, 34 residents and 23 Board certified clinicians fully completed the test. The statistical analysis was performed on these 170 participants. Cronbach's alpha values were 0.67 for the entire group of participants, 0.82 for the fourth-year students, 0.50 for the fifth and sixth-year students, 0.55 for the residents and 0.56 for the practitioners.

Global mean scores for the different groups are shown in Table 2. Differences between the mean scores for the four groups were considered statistically significant ($p < 0.001$). The post-hoc Bonferroni analysis indicated that significant differences were present between fourth-year students and residents ($p < 0.001$), between fourth-year students and board certified practitioners ($p = 0.001$). Differences were present between fourth-year students and the other students (fifth and sixth years), between fifth and sixth-year students and residents, between fifth and sixth-year students and practitioners, but they were not considered significant.

Table 2. Comparison of mean scores by groups according to level of clinical experience in therapeutic care.

Group	Size	Mean score	SD	CI 95%
Students	113	64.06	7.62	62.64–65.48
Fourth-year students	16	58.93	11.42	52.84–65.02
Fifth and Sixth year students	97	64.91	6.50	63.60–66.22
Residents	34	67.65	6.57	65.36–69.94
Board certified practitioners	23	68.47	6.39	65.70–71.24

Notes: Statistical analysis were performed only on the 170 participants out of 251 who fully completed the test. SD: Standard deviation; CI: confidence interval.; One-way analysis of variance: $p < 0.001$. Bonferroni correction analysis: fourth-year students vs. residents, $p < 0.001$; fourth-year students vs. board certified practitioners, $p = 0.001$; entire group of students vs. residents, $p = 0.038$; entire group of students vs. board certified practitioners, $p = 0.027$. Other comparisons between groups were not significant.

There were no differences between residents and practitioners. When we compared mean score for the entire group of students ($N=113$) with other groups, we found a statistically significant difference between students and residents ($p=0.038$) and between students and board certified practitioners ($p=0.027$).

Discussion

This study reports one of the first large-scale utilizations of the Concordance test in France. The major goal of this study was to determine the feasibility of the Concordance test on our learning environment. The assessment of therapeutic decision-making with the Concordance test was challenging for several reasons: First, both French educators and students were not familiar with the format of Concordance test questions; Second, there was a lack of experience of on line evaluation system among French Medical schools; Third, French therapeutics teaching program is linked to numerous medical specialties. It was an arduous task to find 15 practitioners who were not overly sub specialized and therefore could be used as our reference panel. To avoid this potential pitfall, we created a reference panel of Faculty members who had a medical practice basically focused on a multidisciplinary approach of the patient. We opted for this approach because we wanted to measure decision-making skills at an undergraduate level, as assessed by the National ranking Examination.

The use of the Internet tool permitted us to quickly evaluate psychometric properties of the test on a large scale. As regards validity, our results showed that fourth-year students obtained significantly lower results than residents, and than Board certified practitioners. Fourth-year students' performance was lower than fifth and sixth-year students' performance but with no statistical difference. Despite the global correlation of scores with levels of training observed in our study, the construct validity of the concordance test for therapeutics was not straightforward. Nevertheless, in France, therapeutic care is not specifically taught to fourth-year students and it is interesting to note that this less experienced group of participants obtained the lower results. Moreover, reliability coefficients of our concordance test on line appears to be acceptable for a feasibility study.

We acknowledge that our findings may have several limitations. Despite the fact that the individual programs of French medical schools are very similar because the National

Examination is based on common learning objectives, we can not assume that participating students are totally comparable. There are nonetheless slight differences in each medical school teaching organization. An other point is the limited number of fourth-year students and residents. Furthermore, despite the fact that reliability coefficients were acceptable, the low number of test questions may reduce the global quality of the test. Therefore, there is a possible lack of power in the statistical analysis to detect any significance in the difference of the scores between participating groups. The other limitation of our methodology was that only 170 participants out of 251 had entirely completed the test. Based on these results, the fact that participants never received e-mail reminders during the inclusion period should be taken into consideration. Feedback possibilities of our Web site should permit us to focus on improvement in the completion of the Concordance test on line. Another bias was due to the fact that examinees may have possibly consulted resources during examination time or could have exchanged information with others candidates before performing the test. We are currently planning the use of Web-cameras to remove this bias.

Although the present study is limited in several ways, our results are well correlated with other studies on Concordance test which have previously demonstrated the validity and the reliability on the Concordance test in different medical specialties (Brailovsky et al. 2001; Charlin et al. 2002; Charlin & Van der Vleuten 2004; Gagnon et al. 2006; Sibert et al. 2006b; Charlin et al. 2007; Meterissian et al. 2007). Sample size of participants was comparable or greater than those reported in the literature. Furthermore, interesting perspectives could be obtained from our exploratory research. From an educational point of view, it is interesting to note that the Concordance test offers the opportunity of an extensive assessment of decision-making skills in various contexts, which are closer to the reality of clinical practice. We accept that our results, as reported, only demonstrated that the Concordance test was able to discriminate examinees who received therapeutics teaching from those who have not. This reinsures that the Concordance test could be used as a useful complement with other tools of evidence-based knowledge assessment, i.e. patient management problems or rich-context MCQ as used by the National Board of Medical Examination in the United States (De Champlain et al. 2003). These concepts underline the current need to promote this type of Web site. From a psychometric point of view, a web-based assessment with the Concordance

test should have several advantages. Scoring process is objective, standardized, discriminant and can be easily computer-based. Acceptable reliability coefficient was reached within a relatively short examination period (i.e. 0.67 with a 50 min examination duration). Our Web site has provided a potentially cost-effective opportunity to include participants without logistic difficulties traditionally encountered with paper-based evaluation systems. On the other hand, the cost for administration and statistical analysis was the same as a paper-based Concordance test. Moreover, the use of the Internet should permit us to administer the Concordance test on a large scale over a relatively short period of time. Taking into account the very wide audience ($n=5,600$ students) of the French undergraduate ranking examination, our findings warrant further consideration. Nevertheless, before the summative use of Concordance tests in our learning environment, many questions remain. First, we will need to improve Concordance test content validity, with a greater number of clinical cases, in order to closer discriminate among examinees with varying degrees of therapeutics training. An other point of discussion is the composition of the reference panel. For instance, if a Concordance test is taken for a certification purpose for the undergraduate curriculum, should the panel be made up of physicians who teach students or physicians representing the profession? One other major concern is the correlation of the Concordance test with National ranking Examination in order to establish the added value of Concordance test to existing instruments for assessing decision-making skills. Furthermore, the application of a standard setting procedure for the Concordance test remains to be established. Future studies should be directed towards these particular issues in order to facilitate a more accurate approach regarding the utility of this assessment tool as a therapeutic care teaching and assessment strategy.

Conclusion

This study suggests that an on line assessment course of therapeutic competencies with the Concordance test is feasible and can be effectively developed with acceptable reliability. The French medical teaching environment urgently requires further research before including the Concordance test as a part of the National undergraduate ranking Examination. The Concordance test for therapeutics continues to be available on line, which should provide increasing data as regards these major research questions.

Acknowledgements

This research project was funded by a grant from the FMVU consortium. The authors thank Richard Medeiros, Rouen University Hospital Medical Editor, for editing this manuscript.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

Notes on contributors

LOUIS SIBERT, MD, PhD, is an Associate Professor of Urology in the Department of Urology, Rouen University Hospital. He is also a member of the Department of Medical Education and a member of the CISMef (catalog and index of French-speaking Medical sites) team, Rouen University Hospital, Rouen, France.

ROCH GIORGI, MD, PhD, is an Associate Professor of Biostatistics at the Department of Public Health – Medical Informatics, Marseille University Hospital, and a member of LERTIM Lab, EA 3283, Aix-Marseille University, Faculty of Medicine, Marseille, France.

BADISSE DAHAMANA is a Research Engineer, member of the CISMef team, Rouen University Hospital, France.

JEAN DOUCET, MD, PhD, is a Professor of Therapeutics and member of the Department of Medical Education, Rouen University Hospital, Rouen, France. He is also the responsible for the Therapeutics Teaching Program at the Rouen University Hospital Faculty of Medicine.

BERNARD CHARLIN, MD, is a Professor of Surgery. He is also the Director of the Research and Development Unit in Medical Education, Faculty of Medicine, University of Montreal, Montreal, Canada.

STEFAN J. DARMONI, MD, PhD, is a Professor of Medical Informatics, Rouen University Hospital. He is also the Director of CISMef, Rouen University Hospital, and a member of LITIS Lab, EA 4108, Rouen University, France.

Authors' contributions

L. Sibert, S.J. Darmoni and B. Dahamana developed the Web site, which were responsible for data acquisition and revision of this manuscript. R. Giorgi performed the statistical analysis and revised this manuscript. J. Doucet carried out critical revisions. B. Charlin was the inventor of the Concordance test, had the original idea of on line assessment of clinical reasoning with the Concordance test. He was involved in conceiving and planning the study and revised the manuscript. Funding was obtained by L. Sibert and S.J. Darmoni. All authors have read and approved the final manuscript.

References

- Brailovsky C, Charlin B, Beausoleil S, Coté S, van der Vleuten C. 2001. Measurement of clinical reflective capacity early in training as a predictor of clinical reasoning performance at the end of residency: An exploratory study on the script concordance test. *Med Educ* 35:430–436.
- Charlin B, Tardif B, Boshuizen HPA. 2000. Scripts and medical diagnostic knowledge: Theory and applications for clinical reasoning instruction and research. *Acad Med* 75:611–621.
- Charlin B, Desaulniers M, Gagnon R, Blouin D, van der Vleuten C. 2002. Comparison of an aggregate scoring method with a consensus scoring method in a measure of clinical reasoning capacity. *Teach Learn Med* 14:150–156.
- Charlin B, van der Vleuten C. 2004. Standardized assessment of reasoning in contexts of uncertainty: The script concordance approach. *Eval Health Prof* 27:304–319.
- Charlin B, Gagnon R, Sauve E, Coletti M. 2007. Composition of the panel of reference for concordance tests: Do teaching functions have an impact on examinees' ranks and absolute scores? *Med Teach* 29:49–53.
- Cook DA, Dupras DM. 2004. A practical guide to developing effective web-based learning. *J Gen Intern Med* 19:698–707.
- De Champlain AF, Melnick D, Scoles P, Subhiya R, Holtzman K, Swanson D, Angelucci K, McGrenra C, Fournier JP, Benchimol D, et al. 2003. Assessing medical students' clinical sciences knowledge in France: A collaboration between the NBME and a consortium of French medical schools. *Acad Med* 78:509–517.
- French Medical Virtual University. 2008. Web site URL: www.umvf.org [last accessed September 15, 2008].

- Gagnon R, Charlin B, Coletti M, Sauvé E, van der Vleuten C. 2005. Assessment in the context of uncertainty: How many members are needed on the panel of reference of a script concordance test? *Med Educ* 39:284–291.
- Gagnon R, Charlin B, Roy L, St-Martin M, Sauvé E, Boshuizen HPA, van der Vleuten C. 2006. The cognitive validity of the script concordance test: A processing time study. *Teach Learn Med* 18:22–27.
- Lebeux P, Leduff F, Fresnel A, Berland Y, Beuscart R, Burgun A, Brunetaud JM, Chatellier G, Darmoni SJ, Duvauferrier R, et al. 2000. The French Virtual Medical University. *Stud Health Technol Inform* 77:554–562.
- Meterissian S, Zabolotny B, Gagnon R, Charlin B. 2007. Is the script concordance test a valid instrument for assessment of intraoperative decision-making skills? *Am J Surg* 193:248–251.
- Segouin C, Jouquan J, Hodges B, Brechat PH, David S, Maillard D, Schlemmer B, Bertrand D. 2007. Country report: Medical education in France. *Med Educ* 41:295–301.
- Sibert L, Darmoni SJ, Thirion B, Douyere M, Dahamna B, Weber J. 2006a. A web-based teaching resource to prepare for final undergraduate examination: A French pilot study. *Med Teach* 28:158–164.
- Sibert L, Darmoni SJ, Dahamna B, Hellot MF, Weber J, Charlin B. 2006b. On line clinical reasoning assessment with script concordance test in urology: Results of a French pilot study. *BMC Med Educ* 6:45.
- van der Vleuten CPA. 1996. The assessment of professional competence: Developments, research, and practical implications. *Adv Health Sci Educ* 1:41–67.