SHORT REPORT

Usefulness of the script concordance test in dermatology

A.-C. Bursztejn,[†,*] J.-F. Cuny,[†] J.-L. Adam,[‡] L. Sido,[‡] J.-L. Schmutz,[†] J.-D. de Korwin,[‡,§] C. Latarche,[¶] M. Braun,[††,‡‡] A. Barbaud[†]

[†]Dermatology Department, Fournier Hospital, CHU Nancy, Nancy
[‡]Family Medicine Department, Faculty of Medicine, UHP1, Vandoeuvre les Nancy
[§]Internal Medicine Department, Central Hospital, CHU Nancy, Nancy
[¶]Epidemiology and Clinical Evaluation Department, Marin Hospital, CHU Nancy, Nancy
[††]Neuroradiology Department, Central Hospital, CHU Nancy, Nancy
[‡‡]Continuing Education Department, Faculty of Medicine, UHP1, Vandoeuvre les Nancy, France
*Correspondence: A-C Bursztejn. E-mail: ac.bursztejn@chu-nancy.fr

Abstract

Background The script concordance test (SCT) measures clinical data interpretation in the context of uncertainty. To our knowledge, its feasibility has not yet been demonstrated in dermatology.

Objectives This study describes the feasibility, reliability and validity of the SCT for use in dermatology teaching to family medicine residents.

Methods We designed an SCT for dermatology teaching, adapted to the level of family medicine students. The family medicine residents attending a dermatology lecture course and a reference panel of dermatologists underwent the test. Software available on the SCT-dedicated website was used to determine the scores. The Cronbach's α was calculated. The scores were described by means, standard deviation, and minimum and maximum scores. Mann–Whitney tests were used to compare resident and reference panel scores.

Results The test contained 33 vignettes of four questions each. Cronbach's α was 0.80. Mean scores were 75.6 for the reference panel and 65.0 for the residents ($P = 0.0009$).

Conclusions This study confirmed the SCT's capacity to distinguish groups on the basis of experience. This study demonstrated the feasibility and reliability of the SCT in the field of dermatology.

Received: 28 March 2010; Accepted: 26 January 2011

Conflict of interest

The authors do not have any conflict of interest to declare.

Introduction

Although clinical reasoning is a major component of clinical competence in medical education, most evaluating tools do not assess this skill properly. The script concordance test (SCT) was recently developed by Charlin *et al.* to obviate this lack, to assess clinical reasoning in an uncertainty context.[1] The test is based on two important features. First, the 'script' forming the basis of the test is a network of goal-directed knowledge relevant to a given situation that is used to test diagnosis hypotheses, treatments or management options.[1] Scripts usually evolve with training, the most experienced physicians gaining the richest scripts.[1] Second, decision-making in the context of uncertainty is the hallmark of experienced practitioners, whatever their field of expertise is, compared with junior practitioner.[2] For both reasons, the SCT usually distinguish experienced practitioners from junior practitioners.[1]

Many studies have confirmed the utility of the SCT in a variety of medical fields, such as internal medicine, surgery and family medicine.[3–9] As far as we are aware, SCT has not been evaluated in dermatology. The aim of this study was to describe the SCT in dermatology, with particular insights into feasibility, validity and reliability of scores.

Materials and methods

Briefly, the SCT is composed of 'items' including a short vignette and questions that mimic real clinical situations. The clinical situation is described by the vignette. Each item is targeted on a particular field: most frequently, diagnosis, management or treatment. Several options are given, regarding the field of the item, for instance, several diagnosis options for a diagnosis-oriented item. For each option, new relevant information is given, whose impact on the initial option is measured with a five-point Likert scale.

The option, the new information and impact constitute the 'question.' The uncertainty context lies on the vignette (incomplete data), the options (several possible diagnosis or management options for instance) and the relevant new information (for instance, elevated creatinine level and initial prescription of an enhanced CT-scan). Examples are given in Fig. 1.

Score stems from choices on the Likert scale followed by a modal transformation.[10] To establish the answer key, the SCT is taken by a panel of experienced physicians (aggregate scoring method). For any question, the most frequently chosen points will provide a credit of 1 point; the never chosen points will give 0 point, while the other points are given a partial credit (the number of panel members choosing this point divided by the number of panel members selecting the most chosen Likert scale point). In a second step, students take the test. Students' scores are obtained by attributing the afore-mentioned credits to each selected Likert scale point. The total score is the sum of scores obtained for each question normalized to a scale of 100.

Situation: Dermatology course targeted on family medicine residents.

(a)

A man presented with a painful eruption on his face that had been present for 48 h

−2 = rules out or almost ruled out
−1 = less probable
 0 = neither less nor more probable
+1 = more probable
+2 = certain or almost certain

If you were thinking of:	And then you find:	This hypothesis becomes:
An erysipelas	Some grouped vesicles	−2, −1, 0, +1, +2
An impetigo	Pustules	−2, −1, 0, +1, +2
A facial staphylococcus cellulitis	No fever	−2, −1, 0, +1, +2
A zona	A negative direct bacteriological sample	−2, −1, 0, +1, +2

(b)

You find a well limited erythemato-squamous eruption on the elbows that has been present for 10 years. The plaques look like psoriasis. The patient is a 30 y.o. man with no other medical past history.

−2 totally contraindicated
−1 fairly useless or possibly harmful
 0 neither useful nor harmful
+1 useful
+2 indicated or absolutely necessary

If you were considering to prescribe:	And then you find:	With this new information, the relevance of this treatment becomes:
Vitamin D drift (tacalcitol)	Several attacks of articular pain lasting for 2 years	−2, −1, 0, +1, +2
Methotrexate	Occasionnal alcoholism	−2, −1, 0, +1, +2
Class II topical corticosteroid (desonide)	Active smoking estimated at 20 YP	−2, −1, 0, +1, +2
Phototherapy	Recent appearance of a black tumor on his back	−2, −1, 0, +1, +2

(c) A 75 y.o. man has a prurit lasting for one year. This patient takes several drugs: salycilic acid, ramipril, bisoprolol, amlodipine and esomeprazole.

−2 much less useful
−1 less useful
 0 neither more nor less useful
+1 more useful
+2 much more useful

If you were considering to ask:	And then you find:	With this new information, this investigation becomes:
An abdomino-pelvi-thoracic TDM	Asthenia and 4 kg weight loss	−2, −1, 0, +1, +2
A mycologic sample on his back	A whole body papulo-vesicular erythematous eruption	−2, −1, 0, +1, +2
Epicutaneous tests with all drugs taken	All drugs are taken for 6 years without any change	−2, −1, 0, +1, +2
An abdomino-pelvi-thoracic TDM	His wife has the same symptoms lasting for 10 months	−2, −1, 0, +1, +2

Figure 1 Examples of vignettes included in the script concordance test. (a) Diagnostic vignette; (b) Investigation vignette; (c) Treatment vignette.

Development of the SCT

Preparation of the SCT: Clinical situations chosen for the SCT covered dermatological diseases within the scope of general practitioners (GP). These clinical situations were taught during the dermatology course. Clinical situations, vignettes, options and new information were reviewed by the faculty involved in the family medicine curriculum and familiar with the SCT, with particular attention on clarity, required level of competence and appropriateness of uncertainty. Thirty-three vignettes with four independent questions each were prepared. The SCT was elaborated by two dermatologists and two GP, all involved in the family medicine curriculum. The test development followed published recommendations.[10] Hereafter, the dermatological diseases used were: atopic dermatitis and eczema (4 vignettes); psoriasis (5 vignettes); angio-oedema and urticaria (2 vignettes); cutaneous cancers (5 vignettes); ulcers (3 vignettes); infectious diseases (9 vignettes); pruritus (4 vignettes); and facial erythema (1 vignette). After choosing the diseases and corresponding photographs, diagnosis, investigation or treatment, options were provided to write the vignette questions. A PowerPoint presentation was used to administer the SCT. Each slide contained a picture, a short description of the clinical situation, a definition of Likert scale values and a Table including the diagnostic (or investigation or treatment) option, the supplementary information and Likert scale values for respondents to select.

Subjects

A reference panel of privately practising dermatologists was constituted. All belonged to a post-academic society in the Lorraine region. They were asked to answer the SCT anonymously. The SCT was mailed to them and was returned after it was completed. The principle of the SCT was explained over the phone and by means of a short text.

The test was then administrated during a single 1 h session to all family medicine residents attending a dermatology lecture (before the lecture course). All residents were informed about the study's objectives and the principle of the SCT. As the only previous experience, the students had pursued dermatology courses (approximately 21 (hours) lectures) 1–3 years ago.

Scoring method

The students' scores were determined using the automatic calculator developed by the Montréal Medical School, available at http://www.cme.umontreal.ca/tcs. The reference panel scores were manually established.

The calculator computes an item analysis for each question regarding its ability to discriminate between students and its impact on the overall reliability. Thus, each question can be categorized as 'good' (coefficient >0.20), 'fair' (correlation coefficient between 0.10 and 0.20) and 'bad' (correlation coefficient <0.10).

Statistical analysis

All data were entered into Microsoft Excel. The statistical analysis was performed using the SAS software version 9.1 (SAS Institute, San Francisco, CA, USA) at the Epidemiology and Clinical Evaluation Department, CHU Nancy. The internal coherence test coefficient was evaluated with the Cronbach's α test. This statistical index (values between 0 and 1) indicates greater homogeneity as values approach 1. In the evaluating methods, good reliability is indicated when the coefficient is ≥0.80.[11] Scores were determined within two groups: residents and dermatologists. Ideally, two experienced physician groups are requested, one for the score determination and the other for comparison with the students, to appreciate the scores' construct validity. In this research, we were not able to get one experiences group and an expert panel, only 16 dermatologists were available. Each dermatologist's score was determined using a reference score calculated with the 15 other dermatologist answers and excluding its own answers (score).[3] The statistical analysis included the scores of the residents and the experts. Quantitative values were described by the mean, standard deviation and minimum and maximum scores. Confidence intervals of 95% were calculated. Comparison between the student and reference panel scores was undertaken using the Mann–Whiney test because of the non-normal distribution of the data. *P* values <0.05 were considered statistically significant.

Results

The SCT

The test included 22 diagnosis vignettes, seven therapeutic vignettes and four investigation vignettes (details of topics treated in those vignettes are described in the materials and methods section).

Subjects

The test was performed on 35 students. One student answered less than 50% of the test queries and his examination was therefore excluded.
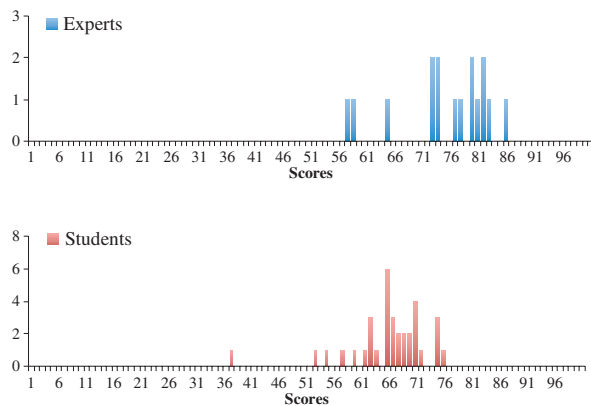
The reference panel was composed of 16 dermatologists from the Lorraine area, with an exclusive private practice.

The scores

The reference panel had a mean SCT score of 75.6 and the residents had a mean score of 65.0 (*P* = 0.0009; Table 1, Fig. 2).

Table 1 Description statistics for reference panel and resident scores based on the script concordance test in dermatology

Scores	Reference panel (*N* = 16)	Residents (*N* = 34)
Mean	74.2	65.0
Standard deviation	8.4	7.3
Minimum	56.5	36.5
Maximum	85.2	75
Confidence interval (95%)	70.9–78.7	63.8–66.3

Figure 2 Comparison of reference panel and resident scores for the script concordance test.

The α Cronbach

The entire SCT had a Cronbach's α of 0.74 and contained 40 bad items (40%), 35 fair items (26.5%) and 57 good items (43.2%). After removing 14 questions with a negative correlation coefficient and a low variance, the Cronbach's α increased to 0.80 and the number of bad items decreased to 27 (22.9%).

Opinion of students and physicians involved in this research

The test was well accepted by experts and students, although some students were surprised by the novelty of the test format.

Most of the practitioners who designed the SCT reported that it was difficult to include the context of uncertainty in the vignettes.

Discussion

This study demonstrates the feasibility of SCT in the field of dermatology, its reliability and validity.

Feasibility

The SCT was written and administrated in a short period, indicating the ease of test construction, despite the uncommon format of the scenarios. Our SCT included more diagnosis questions and only a few investigation items. This was due to the difficulty of maintaining the context of uncertainty in such questions. We easily recruited the panel of experts. The scoring method was not time-consuming, given the available software. Despite the small sample size, Cronbach's α coefficients indicated the reliability of the test.

Scores

Within 1 h of test administration, the SCT could assess eight situations in 33 items and 132 questions. All the situations are commonly found in a GP normal activity. Thus, the attributed scores get correct face and content validities.[4]

As expected, the experienced physicians (dermatologists) had better scores than the junior physicians (residents). This point is of paramount importance because it documents the SCT's ability to distinguish physicians following their clinical experience. This ability confirms the scores' construct validity.[5]

Despite the small sample of residents, the internal value of the SCT (the α Cronbach) was good (0.80), after optimization, reaching the value necessary for high-stake examinations.[10]

In a dermatology setting, the SCT appears as a promising size with attractive intrinsic qualities (validities and reliability of scores) despite a short time for administration.[4] It challenges the multiple choice questions well, but with a constant ability to distinguish practitioners from their level of expertise.[6]

SCT in dermatology

SCT has been widely used in several medical fields with the same results.[6–9,12–14] As far as we are aware, this is the first time it is used in dermatology.

Brazeau-Lamontagne *et al.* gained similar results in imaging, thus evaluating a very close field of expertise.[15] In future, new multimedia tools will incorporate images, sounds or short video, thus enhancing their face validity.[12,14,16]

Particular aspects of the present SCT

The test was designed following published recommendations.[10] The blueprint was designed by dermatologists and GP to prevent the use of overspecialized material.

The composition of the reference panel is crucial.[3,17] Due to the setting of the test (course of dermatology aimed at family medicine residents), we chose exclusive private practice dermatologists rather than hospital or university teaching hospital dermatologists. The choice of experts especially the expected level of competence is deeply important.[17] Sixteen panel members were an adequate number.[10] Increasing the panel members increases the test reliability.[3] More than 20 members induce a slight improvement in reliability, counterbalanced by a lower feasibility.[3] Such an approach is valuable for high-stake examinations in which reliability is a prominent issue.[3,10]

Four questions were associated with each vignette, constituting an item. Research has shown that two to four questions for each vignette were an adequate balance between reliability and feasibility.[18]

The SCT allows several clinical situations to be explored in a short examination. Evaluation of students may thereby be improved, because a good response to one clinical problem does not necessarily correlate with good scores to others (clinical problems).[19] Moreover, these tests closely mimic real medical practice.

SCT and continuous medical education (CME)

As other physicians have, dermatologists have an important CME activity, which is now mandatory in France. SCT has been recently introduced in this setting.[16] SCT with a small number of items and questions is taken by the participants before the beginning of the CME session. Computerized analysis of the scores is performed immediately, and can indicate in which participants are

unfamiliar. Thus, the session manager can focus on these points, enhancing the CME session efficiency.

Conclusion

This study demonstrates the feasibility of the SCT and the validity and reliability of its attributed scores in the field of dermatology. The SCT is the only available test that evaluates clinical reasoning in the context of uncertainty. It should be associated with other formats like multiple choice question or clinical case studies to assess various aspects of clinical competence.

References

1 Charlin B, Roy L, Brailovsky C et al. The Script Concordance test: a tool to assess the reflective clinician. *Teach Learn Med* 2000; **12**: 189–195.

2 Schon DA. *The Reflective Practitioner. How Professionals Think in Action.* Basic Books, New York, 1983.

3 Charlin B, Gagnon R, Pelletier J et al. Assessment of clinical reasoning in the context of uncertainty: the effect of variability within the reference panel. *Med Educ* 2006; **40**: 848–854.

4 Zubair A, Chong S, Khoo E. *Practical Guide to Medical Students Assessment.* World Scientific Publishing Company, New Jersey, 2006.

5 LaDuca A. Validation of professional licensure examinations: profession, theory, test design and construct validity. *Eval Health Professions* 1994; **17**: 178–197.

6 Fournier JP, Thiercelin D, Pulcini C et al. Evaluation du raisonnement clinique en médecine d'urgence: les tests de concordance de script décèlent mieux l'expérience que les questions à choix multiples à contexte riche. *Pédagogie Médicale* 2006; **7**: 20–30.

7 Carriere B, Gagnon R, Charlin B et al. Assessing clinical reasoning in pediatric emergency medicine: validity evidence for a Script Concordance Test. *Ann Emerg Med* 2009; **53**: 647–652.

8 Gibot S, Bollaert PE. Le test de concordance de script comme outil d'évaluation formative en réanimation médicale. *Pédagogie Médicale* 2008; **9**: 7–18.

9 Lambert C, Gagnon R, Nguyen D et al. The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiat Oncol* 2009; **4**: 7.

10 Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. *BMC Med Inform Decis Mak* 2008; **8**: 18.

11 Charlin B, Gagnon R, Sibert L et al. Le test de concordance de script, un instrument d'évaluation du raisonnement clinique. *Pédagogie Médicale* 2002; **3**: 135–144.

12 Caire F, Sol JC, Moreau JJ et al. Auto-évaluation des internes en neurochirurgie par test de concordance de script: processus d'élaboration des tests. *Neurochirurgie* 2004; **50**: 66–72.

13 Meterissian S, Zabolotny B, Gagnon R et al. Is the script concordance test a valid instrument for assessment of intraoperative decision-making skills? *Am J Surg* 2007; **193**: 248–251.

14 Sibert L, Darmoni SJ, Dahamna B et al. On line clinical reasoning assessment with Script Concordance test in urology: results of a French pilot study. *BMC Med Educ* 2006; **6**: 45.

15 Brazeau-Lamontagne L, Charlin B, Gagnon R et al. Measurement of perception and interpretation skills during radiology training: utility of the script concordance approach. *Med Teach* 2004; **26**: 326–332.

16 Kalamarides M, Caire F, Dauger F et al. Créations d'outils d'autoévaluation pour une formation médicale continue en ligne. Modélisation à propos d'une formation. *Neurochirurgie* 2008; **54**: 21–27.

17 Charlin B, Gagnon R, Sauvé E et al. Composition of the panel of reference for concordance tests: do teaching functions have an impact on examinees' rank and absolute score. *Med Teach* 2007; **29**: 49–53.

18 Gagnon R, Charlin B, Lambert C et al. Script concordance testing: more cases or more questions? *Adv Health Sci Educ Theory Pract* 2008; **14**: 367–375.

19 Marie I, Sibert L, Roussel F et al. Le test de concordance de script: un nouvel outil d'évaluation du raisonnement et de la compétence clinique en médecine interne? *Rev Med Interne* 2005; **26**: 501–507.