

# **Script Concordance Test untuk Menilai Kemampuan Penalaran Klinis Dokter: Bagaimana Mengembangkan Butir Soalnya Berdasarkan Standar Kompetensi Dokter Indonesia?**

**Fundhy Sinar Ikrar Prihatanto**

Fakultas Kedokteran Universitas Airlangga  
Surabaya

## **Abstract**

**Background:** Some methods have been developed to assess clinical reasoning skill. It must be understood the principals of each method for the selection if will be applied. In the other hand, it is realized that uncertainty is one of medical context that is poorly measured by other traditional test. The script concordance test (SCT) may be a better method of assessing students, who are taught in the context of uncertainty. SCT could differentiate examinees according to levels of experience in various fields of study. SCT have practical relevance for clinical education.

**Aim:** To give guidance in developing SCT based on Indonesian medical doctor competence

**Method:** Relevant literatures are selected as basis of recommendations of SCT construction.

**Discussion:** Developing SCT needs to consider its: principles as basis of development; format of test item to build a suitable test item, scoring system and panel selection accommodate context of uncertainty, and optimization of test item to enhance SCT.

**Conclusion:** SCT provide investigation on clinical reasoning skill of examinee in the difficult by other objective written methods. Several studies have showing SCT capacity in different levels of experience, and in several field of studies with good results. SCT could be as supplement to assess clinical reasoning skill comprehensively, not replacing other clinical competence assessment method absolutely.

**Keywords:** clinical reasoning skill, script concordance test, panel, scoring system

## **Abstrak**

**Latar Belakang:** Beberapa metode telah dikembangkan untuk menilai keterampilan penalaran klinis. Prinsip-prinsip dari setiap metode untuk pemilihan harus dimengerti sebelum diterapkan. Disisi lain, ketidakpastian merupakan salah satu konteks medis yang sulit diukur dengan tes yang biasa dikerjakan. *Script Concordance Test* (SCT) merupakan metode alternatif untuk menilai keterampilan penalaran klinis dalam konteks ketidakpastian. Dalam berbagai bidang ilmu kedokteran, SCT bisa membedakan berbagai tingkat pengalaman. SCT juga memiliki relevansi praktis untuk digunakan dalam pendidikan klinis.

**Tujuan:** Untuk memberikan panduan dalam mengembangkan SCT berdasarkan kompetensi dokter Indonesia.

**Metode:** Literatur yang relevan dipilih sebagai dasar rekomendasi dalam membuat butir soal SCT.

**Pembahasan:** Diperlukan pertimbangan berikut ini dalam mengembangkan butir soal SCT: prinsip-prinsip dasar, format soal tes untuk membangun butir soal yang sesuai, sistem penilaian dan seleksi panel mengakomodasi konteks ketidakpastian, serta optimalisasi butir soal untuk meningkatkan SCT.

**Kesimpulan:** SCT menilai keterampilan penalaran klinis yang sulit dilakukan dengan metode tertulis obyektif lainnya. Beberapa studi telah menunjukkan kapasitas SCT pada berbagai tingkat pengalaman, dan beberapa studi menunjukkan hasil yang baik. SCT bisa sebagai tambahan untuk menilai keterampilan penalaran klinis secara komprehensif, tetapi tidak menggantikan metode penilaian kompetensi klinis lain.

**Kata kunci:** penalaran keterampilan klinis, *script concordance test*, panel, sistem penilaian

---

Korespondensi: fundhy@fk.unair.ac.id

## PENDAHULUAN

Kemampuan penalaran klinis (*clinical reasoning skills*) mempunyai peran penting dalam praktik klinis. Penalaran klinis dapat didefinisikan sebagai serangkaian proses pembuatan keputusan atau pemecahan masalah yang terlibat dalam suatu masalah kesehatan. Sebagai bagian kompetensi klinis, kemampuan penalaran klinis menjadi tujuan pendidikan kedokteran, pada semua tingkat pendidikan (sarjana, pasca sarjana, profesi dokter, spesialis, sampai pada pendidikan dokter berkelanjutan).<sup>2</sup>

Beberapa metode telah dikembangkan untuk menilai kemampuan penalaran klinis, seperti: *extended matching question* – pertanyaan menjodohkan yang diperluas, pilihan ganda yang didesain untuk menilai kemampuan penalaran klinis, *progress test*, *key feature*, dan diskusi berbasis kasus.<sup>3,4,5,6,7</sup> Tetapi prinsip-prinsip dari setiap metode tersebut harus disadari sebelum metode tersebut digunakan.

Di sisi lain, ketidakpastian telah disadari sebagai konteks dari ilmu kedokteran.<sup>8</sup> Dalam penilaian kompetensi yang diharapkan, metode penilaian seharusnya senyata mungkin dengan konteks praktis. *The script concordance test (SCT)* menawarkan penilaian dalam konteks ketidakpastian yang sulit dinilai dengan metode tes lain.<sup>9</sup> SCT mungkin merupakan metode yang lebih baik dalam menilai mahasiswa, yang diajar oleh sejumlah ahli daripada menggunakan pertanyaan pilihan ganda dengan satu jawaban benar.<sup>10</sup>

SCT dapat membedakan peserta tes sesuai dengan tingkat pengalamannya secara akurat (dari pemula sampai ahli) dalam berbagai bidang ilmu seperti: dalam penguasaan materi diabetes, peresepan obat anti inflamasi non steroid, radiasi onkologi, urologi, ortopedi, kegawatdaruratan, atau neurologi.<sup>11,12,13,14,15,16,17</sup> SCT dapat menjadi alat yang berguna untuk mengukur kemajuan mahasiswa.<sup>11</sup>

Brownel<sup>18</sup> menyatakan bahwa SCT mempunyai relevansi praktis untuk pendidikan klinis dengan beberapa argumen: memperkenalkan suatu alat penilaian berbasis bukti, SCT mempunyai potensi besar sebagai evaluasi

formatif, suatu alat untuk evaluasi sumatif pada berbagai jam 13 tingkat pendidikan berbeda, dan suatu alat *self-assessment* bagi mahasiswa.

## TUJUAN

Tulisan ini ditujukan sebagai penuntun pengembangan SCT, dari prinsip-prinsip dasar ke aspek praktis berdasar kompetensi dokter Indonesia.

## METODE

Literatur yang relevan dipilih untuk melakukan tinjauan. Tinjauan tersebut digunakan sebagai penuntun penyusunan butir soal SCT.

## PEMBAHASAN

### Format Butir Soal

Pengembangan tes berdasar kasus digunakan untuk menilai kemampuan peserta tes. Kasus merupakan skenario pendek yang mengandung ketidakpastian. Beberapa pertanyaan dikembangkan dalam beberapa tinjauan yaitu: promosi dan pencegahan; diagnosis; terapi; dan rehabilitasi.

Sebagai bagian dari penilaian, validitas harus menjadi pertimbangan utama dalam pembuatan butir soal. Keseluruhan tes harus mencakup area penalaran klinis yang diharapkan. Sehingga, *blueprint* menjadi dasar dari pembuatan butir soal. Kompetensi dokter Indonesia (Konsil Kedokteran Indonesia-KKI, 2006)<sup>19</sup> memberikan penyakit yang harus dikuasai dokter, terbagi dalam 4 tingkat. Untuk menilai kompetensi dokter, Tabel berikut ini digunakan sebagai contoh *blueprint* untuk membuat butir soal SCT berdasar penyakit dengan tingkat kompetensi: 3A, 3B, dan 4.

*Blueprint* tersebut menunjukkan jumlah butir soal untuk masing-masing tinjauan. Pada Tabel tersebut digunakan 2 tinjauan sekaligus yaitu sistem organ dan proses normal dan patologis sehingga keseluruhan daftar penyakit dengan tingkat kompetensi yang diharapkan.

Tabel 1. Pemetaan kasus untuk memastikan cakupan sel

			Organ dan sistem										
			Neurobehaviour	Head and Neck	Endocrine and Metabolism	Gastrointestinal, Hepatobiliary and pancreas	Respiratory	Urogenital	Cardiovascular and limfatic	Hemato-immunology	Dermatomusculo-skeletal	Reproductive	TOTAL
			1	2	3	4	5	6	7	8	9	10	
Proses normal dan pathologi	A	Growth, development, and degenerative	1	1	1	1	1	1	1	1	1	1	10
	B	Genetic and congenital	1	2	1	2	2	2	2	1	1	1	15
	C	Infection and immunology	2	2	2	2	2	2	2	2	2	2	20
	D	Neoplasm	1	1	1	1	1	1	1	1	1	1	10
	E	Trauma or accident	2	2	2	2	2	2	2	2	2	2	20
	Total		7	8	7	8	8	8	8	7	7	7	75

**Catatan:**

Angka dalam sel yang diarsir adalah jumlah butir soal sesuai dengan proses normal dan patologi, dari organ dan sistem. Jumlah butir soal sebanyak 75 dengan menggunakan 25 kasus masing-masing yang diikuti 3 butir soal.

Tabel 2. Contoh *blueprint* untuk pembuatan butir soal

No.	Penyakit*	Tinjauan 6 UKDI**			
		Promosi dan pencegahan (20-30%)	Diagnosis (20-30%)	Terapi (20-30%)	Rehabilitasi (10-20%)
1	Bronchitis akut				
2	Asma bronkhial		o		
3	Kolera			o	
4	Disentery basiler			o	
5	Enteritis	o			
6	Alergi makanan		o		
7	Gastritis		o		
8	Gastroenteritis	o			
9	Reflux gastro-esophageal		o		
10	Defisiensi imun - HIV				o
11	Kwashiorkor			o	
12	Marasmus			o	
13	Diabetes tidak tergantung insulin			o	
14	Pneumonia Stafilococcus		o		
15	Tuberkulosis dengan HIV				o
16	Demam tifoid			o	
17	Tuberkulosis paru tanpa komplikasi				o
18	Gastroenteritis viral	o			
19	Defisiensi vitamin			o	
20	Penyakit karena cacing	o			

**Catatan:**

- \* Penyakit tersebut merupakan 20 penyakit dari 142 penyakit dengan tingkat kompetensi 4, kompetensi dokter Indonesia (KKI, 2006)
- \*\* Garis besar komposisi dari materi ujian terdiri dari 7 tinjauan. Tinjauan 6 meliputi: promosi dan pencegahan; diagnosis; terapi; dan rehabilitasi. Komposisi tersebut menentukan jumlah butir soal untuk masing-masing tinjauan (Komite Bersama Uji Kompetensi Dokter Indonesia-KBUKDI, 2007). Pengembangan butir soal selanjutnya berdasarkan diagnosis dan tinjauan yang dimaksudkan tanda (o).

Sebuah kasus mengawali suatu butir soal. Kasus tersebut merupakan deskripsi singkat dalam beberapa kalimat. Deskripsi singkat tersebut harus menampilkan suatu keadaan yang tidak dapat dipastikan pemecahan masalahnya, walaupun oleh seorang ahli.

Gagnon et al.<sup>20</sup> menyarankan bahwa suatu SCT yang reliabel (memiliki koefisien reliabilitas lebih dari 0,80) dapat terdiri dari beberapa pertanyaan per kasus. Sebaiknya sejumlah 20–30 kasus yang diikuti dengan 3–5 pertanyaan per kasus. Penggunaan lebih dari 30 kasus menyebabkan beban kerja yang berlebih bagi pembuat soal dan peserta tes.

**Kotak 1. Rekomendasi pembuatan kasus**

- Suatu kasus dikembangkan berdasar tujuan penilaian (tujuan pembelajaran juga tentunya).
- Setiap kasus mengandung ketidakpastian, beberapa kemungkinan diagnosis and tata laksana.
- Jumlah kasus sebanyak 20-30 buah. Kurang dari 20 mengurangi variasi dan sulit mencakup area penilaian yang diharapkan. Lebih dari 30 kasus membuat beban berlebihan bagi pembuat soal dan peserta ujian.
- Sebuah kasus dapat diikuti 3-5 pertanyaan.
- **Contoh kasus:** Anda adalah seorang dokter yang bertugas di unit gawat darurat. Seorang laki laki 60 tahun datang kepada anda dengan keluhan selama 2 hari. Dia juga menderita batuk yang semakin memberat sejak satu minggu terakhir. Pada pemeriksaan fisik, didapatkan laju pernapasan 24 kali per menit.

Setiap format tes terdiri dari tiga bagian mengikuti setiap kasus yang diberikan, yaitu: **bagian pertama** berisi pilihan diagnosis atau tata laksana yang sesuai; **bagian kedua** menunjukkan suatu temuan klinis, seperti tanda fisik, kondisi tertentu, hasil pencitraan atau pemeriksaan laboratorium; dan **bagian ketiga** sebagai suatu skala Likert untuk memotret keputusan peserta tes.<sup>21</sup>

Peserta tes bertugas menentukan pilihan berdasarkan temuan baru. Skala Likert digunakan untuk menampung konteks ketidakpastian dan dibandingkan dengan jawaban panel ahli. Tiga Tabel berikut ini (**Tabel 2, 3, dan 4**) berisi rekomendasi dan contoh untuk masing-masing tinjauan dengan 2 masalah (kegunaan dan untung-rugi). Contoh yang disampaikan dipilih dan dikembangkan dari diagnosis nomor 1 pada contoh *blueprint* (bronchitis akut) untuk setiap tinjauan dan masalah. Contoh kasus dalam kotak 1 digunakan untuk menyusun butir-butir soal.

Tabel 3.Rekomendasi dan contoh butir soal untuk format diagnosis

Bagian pertama	Bagian kedua	Bagian ketiga				
Bila anda mempertimbangkan diagnosis berikut.....	... dan tersedia informasi baru berikut ini .....	... hipotesis ini menjadi...				
Bronchitis akut	Suara paru vesicular pada kedua sisi	-2	-1	0	1	2

Deskripsi skala Likert: -2: sangat tidak mungkin, -1: tidak mungkin, 0: tidak mendukung ataupun menyingkirkan, +1: mungkin, dan +2: sangat tidak mungkin

Tabel 4.Rekomendasi dan contoh butir soal (masalah kegunaan)

Tindakan	Bagian pertama	Bagian kedua	Bagian ketiga			
Penapisan	Bila anda mempertimbangkan kegunaan tindakan penapisan berikut ini...	... dan tersedia informasi baru berikut ini .....	... anda mempertimbangkan bahwa tindakan penapisan tersebut menjadi.....			
	Rongten dada	Batuk produktif	-2	-1	0	1
Tata laksana	Bila anda mempertimbangkan kegunaan tata laksana berikut ini...	... dan tersedia informasi baru berikut ini .....	... anda mempertimbangkan bahwa tata laksana tersebut menjadi.....			
	Latihan fisik	Gagal jantung kelas fungsional 2	-2	-1	0	1

Deskripsi skala Likert : -2: tidak berguna, -1: sedikit berguna, 0: diantara berguna dan tidak berguna, +1: berguna, dan +2: sangat berguna

Tata laksana dapat dikembangkan berdasar tinjauan promosi dan pencegahan, diagnosis, terapi, dan rehabilitasi. Contoh dikembangkan dari tinjauan rehabilitasi.

Tabel 4. Rekomendasi dan contoh butir soal (masalah untung-rugi)

Tindakan	Bagian pertama	Bagian kedua	Bagian ketiga				
Penapisan	Bila anda mempertimbangkan untung rugi tindakan penapisan berikut ini...	... dan tersedia informasi baru berikut ini .....	...informasi baru tersebut membuat tindakan penapisan menjadi...				
	Bronkioskopi	Pasien yang tidak berpendidikan	-2	-1	0	+1	+2
Tata laksana	Bila anda mempertimbangkan untung rugi tata laksana berikut ini...	... dan tersedia informasi baru berikut ini .....	...selanjutnya anda mempertimbangkan bahwa tata laksana tersebut...				
	Salbutamol	Tekanan darah 150/90 mmHg	-2	-1	0	+1	+2



Untuk butir soal pertama, panel yang terdiri dari 10 orang menjawab: -2 (2 orang), -1 (5 orang), 0 (2 orang), 1 (1 orang), and 2 (0 orang). Pilihan jawaban -1 adalah jawaban dengan jumlah pemilih terbesar. Jumlah pemilih sebesar 5 untuk pilihan jawaban tersebut digunakan sebagai pembagi pada nomor tersebut. Bobot setiap pilihan jawaban pada nomor tersebut secara berurutan adalah: 2/5, 5/5, 2/5, 1/5, and 0/5. Sehingga skor setiap jawaban secara berurutan adalah: 0.4, 1, 0.4, 0.2, and 0.

Selanjutnya, misalkan seorang peserta tes yang menjawab 5 soal tersebut secara berurutan: (nomor 1 sampai 5): -1, 0, -1, 2, dan -1; akan memperoleh skor sebesar:  $1 + 0.4 + 0.125 + 0.25 + 1 = 2.725$  poin dari skor poin maksimal sebesar 5.

### Panel Ahli

Pemilihan panel sangat penting. Untuk menilai kemampuan penalaran klinis pada praktik profesional, Sibert et al.<sup>21</sup> menyarankan bahwa komposisi dari panel harus mempertimbangkan masa depan peserta tes (bergantung pada tujuan penilaian) sehingga pemilihannya harus mengakomodasi beberapa tingkat pengalaman yang berhubungan dengan masalah. Pada tulisan lain, Sibert et al.,<sup>2</sup> mencontohkan pada suatu pendidikan sarjana kedokteran, dalam tujuan pendekatan multidisiplin, anggota panel terdiri dari pengajar, dan ahli penyakit dalam atau praktisi kegawatdaruratan.

Panel akan menentukan bobot dari setiap jawaban. Sehingga, variasi jawaban panel merupakan komponen

Tabel 5. Contoh sistem pemberian skor

Nomor soal		Pilihan jawaban					Contoh dari jawaban peserta tes	
		-2	-1	0	1	2	Pilihan jawaban yang dipilih	Poin yang diperoleh
1	Jawaban panel	2	5*	2	1	0		
	Bobot	2/5	5/5	2/5	1/5	0/5	-1	1
	Skor	0,4	1?	0,4	0,2	0		
2	Jawaban panel	3	5*	2	0	0		
	Bobot	3/5	5/5	2/5	0/5	0/5	0	0.4
	Skor	0,6	1	0,4?	0	0		
3	Jawaban panel	8*	1	1	0	0		
	Bobot	8/8	1/8	1/8	0/8	0/8	-1	0.125
	Skor	1	0,125?	0,125	0	0		
4	Jawaban panel	2	0	0	8*	2		
	Bobot	2/4	0/8	0/8	8/8	2/4	2	0.25
	Skor	0,25?	0	0	1	0,25		
5	Jawaban panel	1	8*	1	0	0		
	Bobot	1/8	8/8	1/8	0/8	0/8	-1	1
	Skor	0,125	1?	0,125	0	0		
Skor maksimal = 5						Skor total yang diperoleh= 2.725		

#### Catatan:

Jumlah anggota panel sebanyak 10 orang

Skor setiap pilihan jawaban = (jumlah panel yang menjawab pilihan jawaban tersebut) / (jumlah pemilih pilihan dengan jumlah pemilih terbesar).

\*jumlah pemilih pilihan dengan jumlah pemilih terbesar, sebagai denominator penentuan pembobotan setiap nomor soal

? skor yang diperoleh untuk setiap nomor soal

Skor total yang diperoleh merupakan penjumlahan perolehan skor dari nomor.

kunci kekuatan pembeda SCT dan dapat membedakan tingkat pengalaman klinis peserta tes.<sup>23</sup> Bahkan seorang dokter yang pengajar dapat menjadi anggota panel, tetapi mereka perlu dilatih karena tes ini belum familiar secara umum.<sup>24</sup>

Pertimbangan selanjutnya adalah jumlah anggota panel. Gagnon, et al.<sup>25</sup> menunjukkan bahwa jumlah di atas 10 orang berhubungan dengan reliabilitas yang dapat diterima dan korelasi yang baik, tetapi lebih dari 20 orang anggota tidak disarankan.

#### Kotak2. Pertimbangan pemilihan panel

- Pemilihan panel bertujuan untuk mencapai tujuan penilaian (tujuan pembelajaran juga tentunya).
- Konteks ketidakpastian tercermin dari variasi jawaban panel.
- Pilihlah panel yang berhubungan kompetensi peserta tes yang diharapkan.
- Ingatlah, mengumpulkan sejumlah panel (dan untuk menjawab soal) adalah hal sulit.
- Sejumlah 10 sampai 20 orang anggota panel adalah cukup. Kurang dari 10 orang menghasilkan variasi yang rendah, sedangkan lebih dari 20 orang tidak menghasilkan peningkatan yang bermakna.
- Contoh anggota panel: untuk menilai kompetensi dokter dipilih panel yang terdiri dari 2 dokter, 2 ahli penyakit dalam, seorang ahli paru, seorang ahli anestesi, 2 ahli bedah, seorang ahli jantung, 2 ahli anak, 2 ahli kandungan, dan seorang ahli saraf.

#### Optimalisasi

Dubois menyatakan bahwa butir soal sebaiknya diperkaya dengan pemaparan peserta tes terhadap masalah klinis senyata mungkin. Pencitraan radiografi dan *scanning* dapat digunakan sebagai data tambahan.<sup>26</sup> Menurut pendapat penulis, modalitas menggunakan indera lain seperti suara (sangat mungkin bila tes menggunakan komputer), bau (bila memungkinkan), atau gambar lain (EKG atau CTG) atau gambar klinis, dapat pula digunakan sehingga kasus yang disajikan mendekati seting nyata.

SCT dapat dilaksanakan secara *paper based*, *computer based* atau melalui sistem *online*.<sup>21,26</sup> Penggunaan komputer memberikan kemudahan dalam mengelola tes obyektif seperti SCT seperti kemudahan untuk pengumpulan nilai tanpa adanya *human error*, terutama bila peserta tes berjumlah besar. Dengan bantuan komputer sangat mudah mengelola data dan melaksanakan tes, bahkan bila dilaksanakan pada tempat yang berjauhan. Tetapi, disamping keuntungannya, pemilihan komputer dan/atau internet ke dalam penilaian membutuhkan persiapan sumber daya yang baik (seperti *hardware* pendukung, pengembangan *software*, dan sumber daya manusia yang terlatih).

#### KESIMPULAN

SCT menilai kemampuan penalaran klinis peserta tes yang sulit dinilai dengan menggunakan metode penilaian tertulis obyektif lain.<sup>9</sup> Beberapa penelitian menunjukkan kemampuan SCT membedakan tingkat pengalaman dan dalam berbagai bidang ilmu.

Seperti yang disadari bersama bahwa tidak ada metode penilaian yang sempurna, SCT tidak dapat menggantikan metode penilaian kompetensi klinis lain secara keseluruhan. SCT dapat digunakan sebagai pelengkap untuk menilai kemampuan penalaran klinis secara komprehensif.

Format SCT memungkinkan penilaian terhadap segi-segi penalaran klinis yang biasanya dikesampingkan dari penilaian yang tradisional tetapi sering dihadapi dalam praktik klinis sehari-hari: penalaran dalam situasi dengan jawaban yang jelas dan benar sulit didapatkan.

#### UCAPAN TERIMA KASIH

Penulis berterimakasih kepada dr. Gandes RR, Ph.D, Prof. Amitya, Prof. Rita, dan dr. Ova E, Ph.D; atas bimbingan beliau dalam mempelajari *assessment* dalam pendidikan kedokteran.

#### DAFTAR PUSTAKA

1. Vertue FM, Haig BD. An abductive perspective on clinical reasoning and case formulation. *Journal of Clinical Psychology*. 2008; 64(9): 1046–68.
2. Sibert L, Giorgi R, Dahamna B, Doucet J, Charlin B, Darmoni SJ. Is a web-based concordance test feasible to assess therapeutic decision-making skills in a

- French context? *Medical Teacher*. 2009;31: e162-e168.
3. Beullens J, Struyf E, van Damme B. Do extended matching multiple-choice questions measure clinical reasoning? *Medical Education*. 2005;39: 410-7.
  4. Elstein AS. Beyond multiple-choice question and essay: the need for a new way to assess clinical competence. *Academic Medicine*. 1993; 68(4):244-9.
  5. Boshuizen HPA, van der Vleuten C, Schmidt HG, Maciels-Bongaerts M. Measuring knowledge and clinical reasoning skills in a problem-based curriculum. *Medical Education*. 1997;31: 115-21.
  6. Pinnock R, Jones A. An undergraduate paediatric curriculum based on clinical presentations and 'key features'. *Journal of Paediatrics and Child Health*. 2008;44:661-4.
  7. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE guide no. 3. *Medical Teacher*. 2007;29: 855-71.
  8. De la Fish, Cole. *Curriculum Development*, Open University Press, 2005.
  9. Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. *BMC Medical Informatics and Decision Making*. 2008; 8:18.
  10. Duggan P. Development of a script concordance test using an electronic voting system. *Ergo*. 2007;1:35-41.
  11. Khonputsa P, Besinque K, Fisher D, and Gong WC. Use of script concordance test to assess pharmaceutical diabetic care: a pilot study in Thailand. *Medical Teacher*. 2006;28(6): 570-3.
  12. Labelle M, Beaulieu M, Paquette D, Fournier C, Bessette L, Choquette D, Rahme E, Thivierge RL. An integrated approach to improving appropriate use of anti-inflammatory medication in the treatment of osteoarthritis in Que'bec (Canada): the CURATA model. *Medical Teacher*. 2004;26(5): 463-70.
  13. Lambert C, Gagnon R, Nguyen D, Charlin B. The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiation Oncology*. 2009;4:7.
  14. Sibert L, Darmoni SJ, Dahamna B, Hellot MF, Weber J, Charlin B. On line clinical reasoning assessment with script concordance test in urology: results of a french pilot study. *BMC Medical Education*. 2006; 6:45.
  15. Cohen LJ, Fitzgerald SG, Lane S, Boninger ML, Minkel J, McCue M. Validation of the seating and mobility script concordance test. *Assistive Technology*. 2009;21:47-56.
  16. Humbert A. Assessing the clinical reasoning skills of emergency medicine clerkship students using a script concordance test. *Acad. Emerg. Med*. 2008; 15(5): S231.
  17. Lubarsky S, Chalk C, Kazitani D, Gagnon R, Charlin B. The script concordance test: a new tool assessing clinical judgment in neurology. *Can. J. Neurol. Sci*. 2009;36: 326-31.
  18. Brownel AKW. The script concordance tests. *Can. J. Neurol. Sci*. 2009;36:272-3.
  19. *Konsil Kedokteran Indonesia. Standar kompetensi dokter*. Jakarta: KKI, 2006.
  20. Gagnon R, Charlin B, Lambert C, Carrie're B, van der Vleuten C. Script concordance testing: more cases or more questions? *Adv in Health Sci. Educ*. 2009; 14:367-75.
  21. Sibert L, Darmoni SJ, Dahamna B, Weber J, Charlin B. Online clinical reasoning assessment with the script concordance test: a feasibility study. *BMC Medical Informatics and Decision Making*. 2005; 5:18.
  22. Charlin B, Desaulniers M, Gagnon R, Blouin D, van der Vleuten C. Comparison of an aggregate scoring method with a consensus scoring method in a measure of clinical reasoning capacity. *Teaching and Learning in Medicine*. 2002;14(3): 150-6.
  23. Charlin B, Gagnon R, Pelletier J, Coletti M, Abi-Rizk G, Nasr C. Sauve E, van der Vleuten C. Assessment of clinical reasoning in the context of uncertainty: the effect of variability within the reference panel. *Medical Education*. 2006;40: 848-54.
  24. Charlin B, Gagnon C, Sauvé E, Coletti M. Composition of the panel of reference for concordance tests: do teaching functions have an impact on examinees' ranks and absolute scores? *Medical Teacher*. 2007;29(1):49-53.
  25. Gagnon R, Charlin B, Coletti M, Sauve E, van der Vleuten C. Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? *Medical Education*. 2005; 39: 284-91.
  26. Dubois JM, Michenaud C, Isidori P. A new way to assess medical competencies: the script concordance test (SCT) on line. *Current Developments in Technology-Assisted Education*. 2006; 1143-7.