

# **STUDY CONCURRENT VALIDITY ANTARA SCRIPT CONCORDANCE TEST (SCT) DAN MULTIPLE CHOICE QUESTIONS (MCQ) DALAM MENILAI PENALARAN KLINIK MAHASISWA KEDOKTERAN**

Yeny Dyah Cahyaningrum\*, Gandes Retno Rahayu\*\*, Efrayim Suryadi\*\*

\* Fakultas Kedokteran Universitas Islam Indonesia

\*\* Departemen Pendidikan Kedokteran, Fakultas Kedokteran Universitas Gadjah Mada

## **ABSTRACT**

**Background:** In the teaching process at medical faculty, there are many measuring devices have been used to evaluate the clinical reasoning ability. Each measuring device has the advantage and disadvantage. MCQ and SCT are the clinical reasoning measuring device. In the previous study, both methods have been validated. SCT is a more recent method to measure clinical reasoning compared to MCQ. This study aims to evaluate the concurrent validity between MCQ and SCT.

**Method:** Forty-three students who had already passed the Paediatric clinical rotation were asked to complete 100 items of MCQ. The next day, they did 100 items of SCT. The concurrent validity is measured using correlation analysis.

**Results:** Concurrent validity between MCQ and SCT did not show strong correlation (coefficient of rank spearman correlation= 0.291 with the probability = 0.058).

**Conclusion:** Concurrent validity between MCQ and SCT do not have strong correlation in evaluating the clinical reasoning. SCT is used as complement to evaluation of clinical reasoning in order to be more specific and close to real setting.

**Keywords:** concurrent validity, script concordance test, multiple choice question, clinical reasoning

## **ABSTRAK**

**Latar belakang:** Dalam proses pembelajaran di Fakultas Kedokteran, banyak alat ukur yang telah digunakan untuk menilai kemampuan penalaran klinik. Masing masing alat ukur ini mempunyai kelebihan dan kekurangan. MCQ dan SCT merupakan alat ukur penalaran klinik. Pada penelitian terdahulu telah dilakukan validasi kedua alat ukur tersebut. SCT merupakan alat ukur penalaran klinik yang relatif baru dibandingkan dengan MCQ. Penelitian ini bertujuan untuk menilai *concurrent validity* antara MCQ dan SCT.

**Metode:** Empat puluh tiga mahasiswa yang telah selesai mengikuti stase Ilmu Kesehatan Anak diminta untuk mengerjakan 100 soal MCQ. Sehari sesudahnya mereka kembali untuk mengerjakan 100 soal SCT. Analisis *concurrent validity* dengan uji korelasi.

**Hasil:** *Concurrent validity* antara MCQ dan SCT mempunyai hubungan yang tidak kuat (koefisien korelasi *rank spearman* = 0.291 dengan probabilitas = 0.058).

**Kesimpulan:** *Concurrent validity* antara SCT dengan MCQ tidak memiliki hubungan yang kuat pada penilaian penalaran klinik. SCT digunakan untuk melengkapi penilaian penalaran klinik agar lebih spesifik dan mendekati pada *real setting*.

**Kata kunci:** *concurrent validity*, *script concordance test*, *multiple choice questions*, penalaran klinik

---

korespondensi: yenydyah@yahoo.com

## PENDAHULUAN

Penalaran klinik merupakan sebuah pemikiran dan pengambilan keputusan masalah pada praktik profesional untuk memandu praktik.<sup>1,2</sup> Kemampuan untuk mendiagnosis efektif dan tepat membutuhkan pengetahuan klinik dan kemampuan penalaran klinik yang baik.<sup>3</sup> Penalaran klinik dinilai menggunakan situasi profesional yang autentik sebagai rangsangannya, bentuknya dapat merupakan sebuah tulisan, verbal, atau mendemonstrasikan situasi dalam praktik.<sup>4</sup>

MCQ merupakan salah satu bentuk penilaian yang simpel dan tradisional untuk menunjukkan kemampuan penalaran klinik.<sup>4,5</sup> MCQ ini digunakan untuk menilai mahasiswa menggunakan kemampuannya untuk menyelesaikan kasus yang permasalahannya pasti, solusinya jelas.<sup>5</sup> Selain itu, MCQ dapat mengukur *factual knowledge* dari mahasiswa.<sup>6</sup> Akan tetapi dalam implementasinya, MCQ mempunyai beberapa kekurangan. MCQ kesulitan untuk menilai *ill defined problem*- masalah yang tidak pasti, sebagaimana ada dalam dunia kedokteran, bahwa situasi-situasi yang mirip dalam kedokteran belum tentu memuat data dan urutan yang pasti sama.<sup>5</sup>

Sebuah alat ukur penilaian tertulis lain yang dapat menilai penalaran klinik mahasiswa adalah *Script Concordance Test (SCT)*. Alat ukur ini dirancang untuk mengukur organisasi pengetahuan yang digunakan untuk menginterpretasikan data dalam penalaran klinik.<sup>9</sup> SCT juga menggunakan kasus yang autentik ada di klinik dalam praktik profesional. Kasus yang digunakan dalam alat ukur ini adalah *ill defined problem*, yang masing-masing ahli tidak selalu mempunyai pendekatan/cara berpikir yang sama dan prinsip-prinsip yang sama tentangnya, sehingga kasus yang ada ini digunakan untuk membandingkan keputusan (penilaian) mahasiswa dibandingkan dengan klinisi. Nilai yang didapat pada SCT akan membandingkan derajat urutan penampilan mahasiswa dibandingkan dengan *panel review*.<sup>10</sup> SCT menilai kompetensi klinik dalam situasi yang autentik dan membutuhkan kemampuan kognitif yang cukup baik.<sup>11</sup>

Dalam alat ukur ini, kemampuan penalaran klinik mahasiswa didasarkan pada kemampuan pembentukan *illness script* yang dibentuk dari latihan/

kemampuan individu dan pengalaman klinik yang pernah dialami.<sup>4</sup> Pada perjalanan profesional seorang dokter, keterampilan dan pengalaman yang dimiliki akan berbeda karena terdapat proses elaborasi pengalaman dan pengetahuan terkait dengan hal-hal rutin yang dilakukan.<sup>12</sup> Kemampuan mendiagnosis secara efektif dan akurat membutuhkan pengetahuan klinik dan keterampilan penalaran klinik yang memadai.<sup>3</sup>

Beberapa penelitian terdahulu yang sudah menggunakan SCT sebagai alat ukur penalaran klinik menyatakan bahwa SCT memiliki kelebihan, yaitu dapat mengukur kemampuan mahasiswa dalam menyelesaikan diagnosis klinik yang terkait dengan ketidakpastian,<sup>13</sup> sehingga yang dinilai pada SCT adalah tidak hanya sekedar *factual recall* saja. SCT disusun untuk dapat mengukur hal yang spesifik dan penting dalam penalaran klinik, yaitu interpretasi data klinis.<sup>13</sup> Dalam mengerjakan SCT mahasiswa membuat keputusan dalam masalah terkait dengan penyakit seperti pada praktik profesional yang akan mereka lalui.

Salah satu kelemahan SCT adalah validitas interpretasinya belum ditetapkan dalam alur yang baku.<sup>14</sup> Sistem *scoring* dalam SCT yang melibatkan ahli merupakan sistem yang cukup kompleks dalam implementasinya. Kelemahan dari SCT yang lain adalah banyaknya jumlah ahli yang dibutuhkan untuk *standard setting*.<sup>11</sup> Selain itu pembuatan soal yang sesuai dan cocok dengan situasi klinik ini merupakan hal yang sulit dan membutuhkan waktu yang panjang.

Validitas dan reliabilitas SCT dalam menilai penalaran klinik telah lebih jauh dilihat dalam beberapa bidang ilmu seperti radiologi, neurologi, radio-onkologi, *surgery*, atau kegawatdaruratan.<sup>13</sup> Dari beberapa penelitian terdahulu telah terbukti bahwa SCT terbukti reliabel dengan *cronbach alpha coefficient* sampai pada nilai 0,75 dan menunjukkan adanya *construct validity*, bahwa item-item dalam SCT dapat membedakan kemampuan berdasarkan pengalaman klinis.<sup>13</sup>

Dengan mempertimbangkan aspek penilaian yang ada pada SCT dan MCQ dan bahwa SCT dan MCQ mempunyai beberapa kelebihan dan

kekurangan, maka akan diteliti *concurrent validity* kedua alat ukur tersebut. Penelitian ini merupakan pengembangan dari penelitian terdahulu yang telah dilakukan tentang SCT. Tujuan penelitian ini adalah mengetahui *concurrent validity* antara score SCT dan MCQ yang diberikan pada mahasiswa di akhir stase Ilmu Kesehatan Anak untuk menilai penalaran klinik.

## METODE

Pada penelitian ini diberikan soal ujian stase Ilmu Kesehatan Anak dalam bentuk MCQ. Sehari sesudah mengerjakan MCQ, mahasiswa akan dikumpulkan kembali untuk mengisi SCT (sebagai sebuah alat ukur penalaran klinik). Subyek penelitian terdiri dari 43 orang mahasiswa yang telah selesai mengikuti stase Ilmu Kesehatan Anak dalam periode penelitian. Penelitian ini melalui beberapa tahapan

mulai dari pembuatan *blue print* soal SCT dan MCQ, pembuatan soal, sampai dengan tahapan *concurrent validity* dan pemberian kuesioner pada mahasiswa.

## HASIL DAN PEMBAHASAN

### Pembuatan *blue print assessment* dan perancangan soal MCQ dan SCT

Pembuatan *blue print assessment* ini dilakukan dengan mengacu pada buku panduan pendidikan klinik stase Ilmu Kesehatan Anak FK UII. Dalam *blue print* ini ditentukan 20 diagnosis yang paling sering terdapat di lapangan terkait dengan masalah pada anak (tabel 1). Keduapuluh jenis diagnosis yang dipilih sekaligus digunakan untuk set soal MCQ agar kesepadanan dapat dijaga. Fokus *blue print* terdiri dari *diagnostic knowledge* (30-60%), *treatment knowledge* (30-60%), dan *investigation knowledge* (5-15%).

Tabel 1. Contoh Blue Print Script Concordance Test dan MCQ

No	Penyakit	Tujuan Belajar	Tingkat kompetensi	Fokus
1	Common Cold	Mampu membuat diagnosis klinik berdasarkan pemeriksaan fisik dan pemeriksaan fisik tambahan yang diminta serta dapat menangani problem secara mandiri terkait penyaki Common Cold pada anak	4	<i>Diagnostic knowledge</i>  <i>Treatment knowledge</i>
2	Faringitis	Mampu membuat diagnosis klinik berdasarkan pemeriksaan fisik dan pemeriksaan fisik tambahan yang diminta serta dapat menangani problem secara mandiri terkait penyaki t Faringitis pada anak	4	<i>Diagnostic knowledge</i>  <i>Treatment knowledge</i>
3	Bronkitis	Mampu membuat diagnosis klinik berdasarkan pemeriksaan fisik dan pemeriksaan fisik tambahan yang diminta serta dapat menangani problem secara mandiri terkait penyakit Bronkitis pada anak	4	<i>Diagnostic knowledge</i>
4	Pneumonia	Mampu membuat diagnosis klinik berdasarkan pemeriksaan fisik dan pemeriksaan fisik tambahan yang diminta serta dapat memutuskan dan memberi terapi pendahuluan dan merujuk (kasus gawat darurat) terkait pneumonia pada anak	3B	<i>Diagnostic knowledge</i>

Sumber: data primer

Perancangan set soal MCQ dan SCT dilakukan oleh peneliti dan didampingi oleh pakar Ilmu Kesehatan Anak. Perancangan MCQ dilakukan dengan mengacu pada kaidah pembuatan soal yang benar. MCQ terdiri dari *stem soal*, *lead in*, dan *option jawaban*. Option jawaban yang disusun ini homogen sesuai dengan *lead in* pada soal.

Perancangan soal SCT dilakukan dengan menuliskan *stem soal* dengan kasus yang mendekati keadaan

klinis yang sebenarnya. Dalam perancangan SCT kasus merupakan kondisi yang kejadiannya tidak bisa didiagnosis dengan pasti untuk sebuah diagnosis (untuk *diagnostic knowledge*). Hal ini juga dilakukan dalam perancangan SCT pada *investigation knowledge* dan *treatment knowledge*. Tabel 2 menunjukkan contoh perbedaan rancangan MCQ dan SCT pada pemeriksaan penunjang Typhoid.

Tabel 2. Perbandingan Rancangan MCQ dan SCT pada Investigation Typhoid

Tujuan Belajar	MCQ	SCT																																																
Mampu membuat diagnosis klinik dan diagnosis banding berdasarkan pemeriksaan fisik dan pemeriksaan fisik tambahan yang diminta serta dapat menangani problem secara mandiri terkait penyakit Typhoid pada anak	Seorang anak laki-laki berusia 10 tahun BB: 20 kg dibawa orang tuanya ke UGD karena demam sejak 1 minggu yang lalu. Demam dirasakan terutama malam hari, dan semakin lama demam semakin meningkat. Keluhan disertai mual, muntah, dan penurunan nafsu makan. Pada pemeriksaan fisik, dokter menemukan adanya bradikardi relatif, <i>typhoid tongue</i> , dan hepatosplenomegali. Apakah pemeriksaan penunjang gold standar untuk menegakkan diagnosis pasien? A. Pemeriksaan darah rutin B. Tes fungsi hati C. Widal <b>D. Gall Culture</b> E. Tubex TF Apakah obat lini pertama yang dapat diberikan pada pasien tersebut? A. Injeksi Ceftriaxone 1 - 2 gram/hari B. Cotrimoxazole 2 x 480 mg peroral C. <b>Chloramphenicol 4 x 500 mg peroral</b> D. Azithromycin 1 gram/hari peroral E. Amoxicillin 3 x 1 gram peroral	Seorang anak laki-laki berusia 5 tahun dibawa orang tuanya ke rumah sakit dengan keluhan demam sejak seminggu yang lalu. Sudah diberi penurun panas. Setelah diberi penurun panas, panas turun, tapi tak lama kemudian panas naik lagi. Demam lebih sering terjadi pada malam hari. Intake makan dan minum dalam seminggu terakhir ini sangat menurun.																																																
		<table border="1"> <thead> <tr> <th>No Soal</th> <th>Bila Anda mempertimbangkan pemeriksaan penunjang</th> <th>Dan tersedia informasi baru berikut ini.....</th> <th colspan="5">Maka pemeriksaan yang Anda pikirkan menjadi.....</th> </tr> </thead> <tbody> <tr> <td>43</td> <td>Darah Rutin</td> <td>Dalam seminggu ini pasien BAB baru 1 kali.</td> <td>-2</td> <td>-1</td> <td>0</td> <td>+1</td> <td>+2</td> </tr> <tr> <td>44</td> <td>Tubex TF</td> <td>Pemeriksaan darah didapatkan leukositosis</td> <td>-2</td> <td>-1</td> <td>0</td> <td>+1</td> <td>+2</td> </tr> <tr> <td>45</td> <td>Widal</td> <td>Ig M anti thypoid (+)</td> <td>-2</td> <td>-1</td> <td>0</td> <td>+1</td> <td>+2</td> </tr> <tr> <td>46</td> <td>Urin rutin</td> <td>Tidak ada gangguan BAK</td> <td>-2</td> <td>-1</td> <td>0</td> <td>+1</td> <td>+2</td> </tr> <tr> <td>47</td> <td>Fungsi Hepar</td> <td>Perut terasa penuh</td> <td>-2</td> <td>-1</td> <td>0</td> <td>+1</td> <td>+2</td> </tr> </tbody> </table>	No Soal	Bila Anda mempertimbangkan pemeriksaan penunjang	Dan tersedia informasi baru berikut ini.....	Maka pemeriksaan yang Anda pikirkan menjadi.....					43	Darah Rutin	Dalam seminggu ini pasien BAB baru 1 kali.	-2	-1	0	+1	+2	44	Tubex TF	Pemeriksaan darah didapatkan leukositosis	-2	-1	0	+1	+2	45	Widal	Ig M anti thypoid (+)	-2	-1	0	+1	+2	46	Urin rutin	Tidak ada gangguan BAK	-2	-1	0	+1	+2	47	Fungsi Hepar	Perut terasa penuh	-2	-1	0	+1	+2
No Soal	Bila Anda mempertimbangkan pemeriksaan penunjang	Dan tersedia informasi baru berikut ini.....	Maka pemeriksaan yang Anda pikirkan menjadi.....																																															
43	Darah Rutin	Dalam seminggu ini pasien BAB baru 1 kali.	-2	-1	0	+1	+2																																											
44	Tubex TF	Pemeriksaan darah didapatkan leukositosis	-2	-1	0	+1	+2																																											
45	Widal	Ig M anti thypoid (+)	-2	-1	0	+1	+2																																											
46	Urin rutin	Tidak ada gangguan BAK	-2	-1	0	+1	+2																																											
47	Fungsi Hepar	Perut terasa penuh	-2	-1	0	+1	+2																																											
		Keterangan: -2: kontraindikasi; -1: tidak berguna; 0: sedikit bermanfaat; +1: bermanfaat; +2: sangat penting																																																

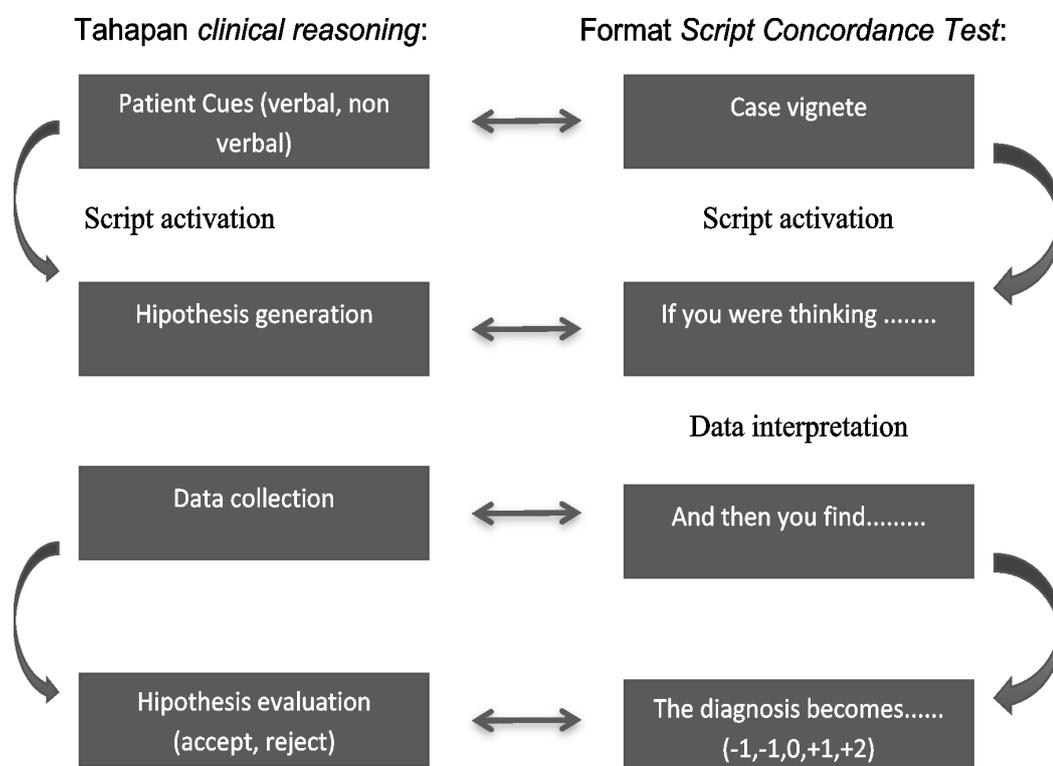
Perancangan *blue print* dilanjutkan dengan pengembangan set soal MCQ dan SCT. Dalam penyusunan MCQ diawali penulisan *stem* soal yang menyediakan stimulus untuk direspon dan menampilkan permasalahan yang seharusnya dihadapi.<sup>18</sup> *Stem* soal berisi masalah atau pertanyaan yang akan dijawab dengan jelas.<sup>15</sup> *Stem* soal yang disusun dalam MCQ ini mengacu pada diagnosis dan *blue print* yang telah disusun.

Kunci jawaban pada MCQ diberikan pada 1 pilihan jawaban saja. Pilihan jawaban yang benar diletakkan acak pada masing-masing soal, untuk menghindari "*placement bias*". Dalam penulisan pilihan jawaban juga memperhatikan komposisi distraktor yang dituliskan di sekitar pilihan jawaban. Pilihan jawaban yang disusun sebaiknya homogen, mirip baik dari sisi konten maupun panjangnya tulisan. Distraktor yang baik adalah pilihan jawaban yang mungkin dipilih yang oleh mahasiswa yang kurang paham, tetapi dipilih oleh mahasiswa yang benar-benar paham masalah tersebut.<sup>18</sup>

Perancangan soal SCT dimulai juga dengan melihat item *blue print* yang tersedia. Peneliti merancang kasus yang merupakan *ill defined problems*. Rancangan soal

pada SCT diharapkan dapat menilai pembentukan *illness script* pada mahasiswa dan melihat konkordansi dengan jawaban panel pakar. Teori *script* ini didasarkan pada teori psikologi kognitif, tentang mekanisme suatu informasi tersimpan dalam memori dan digunakan untuk menginterpretasikan sesuatu.<sup>19</sup> Dalam ilmu kesehatan, teori *Script* ini ditunjukkan dengan hubungan antara keadaan klinik dan patofisiologi penyakit yang menunjukkan kategori diagnostik.<sup>19</sup>

Penilaian soal SCT berbeda dengan proses penilaian pada umumnya. Soal SCT yang dirancang ini memungkinkan semua pilihan jawaban benar dan dapat diterima. Hal ini akan membuat panel menjawab sesuai dengan pengalaman klinis yang pernah didapatkan.<sup>19</sup> Pengalaman klinis dari seorang pakar akan berkembang dari waktu ke waktu. Hal ini sesuai dengan hasil penelitian lain yang menyebutkan bahwa *illness script* akan berkembang dimulai dari mahasiswa memperoleh paparan kondisi klinis, dan disempurnakan dengan pengalaman klinis sesudahnya.<sup>19</sup> Proses mahasiswa mengerjakan soal SCT ini mirip dengan tahapan penalaran klinis yang terjadi pada mahasiswa (gambar 1).



Gambar 1. Hubungan antara tahap penalaran klinis dengan format SCT<sup>16</sup>

### **Content validity dan uji MCQ dan SCT**

Analisis *content validity* dilakukan menggunakan *Lawshe content validity ratio*. Dalam melakukan analisis ini dilakukan oleh 5 orang pakar. Pakar ini merupakan dosen pembimbing klinik dalam pendidikan kedokteran. Hasil *content validity* SCT yang dilakukan dengan *Lawshe content validity rasio* didapatkan nilai sama dengan 1 ( $>0,99$ ) yang menunjukkan bahwa soal SCT memiliki *content validity* yang baik.

Untuk soal MCQ, hasil yang didapatkan adalah soal MCQ dengan *Lawshe's content validity ratio*  $>0,99$  adalah 96 soal. Terdapat soal dengan *Lawshe's content validity ratio* 0,5-0,99 sejumlah 4 soal. Beberapa soal dengan *Lawshe's content validity ratio* 0,5-0,99 mendapatkan beberapa masukan dari pakar, dan selanjutnya dilakukan perbaikan oleh peneliti.

Analisis *content validity* ini dilakukan dengan melihat kesesuaian antara *learning objective* dan kasus yang diberikan. Dalam analisis ini pakar menentukan kategori *essential* atau *not necessary*. Penelitian lain menunjukkan bahwa 50% persetujuan pakar dapat mengatakan bahwa instrument tersebut baik dari sisi *content validity*.<sup>14</sup>

Uji *clarity* dilakukan dengan memberikan set soal SCT kepada mahasiswa yang sudah selesai melakukan stase. Selain itu diharapkan penyakit yang diberikan dalam SCT sudah mereka dapatkan dalam pendidikan klinik yang sudah dialami.

### **Panel scoring SCT**

*Panel scoring* dilakukan oleh 10 pakar Ilmu Kesehatan Anak. Hal ini dilakukan dengan pertimbangan presisi yang didapatkan pada penggunaan 10 *panel scoring* (0,70).<sup>15</sup> Pakar yang melakukan skoring adalah yang memenuhi persyaratan penelitian.

Hasil dari *panel scoring* ini, tidak semua *option* jawaban dalam masing-masing nomor dipilih oleh pakar sebagai jawaban. Dari 100 soal SCT yang diberikan, 36 soal (36%) yang *option*-nya merata dipilih oleh

pakar. Dua puluh satu nomor soal (21%) para pakar lebih cenderung menjawab pada *option* +1, dan +2. Dua nomor soal (2%) dimana para pakar cenderung menjawab *option* -1 dan -2.

### **Reliabilitas SCT dan MCQ**

Evaluasi reliabilitas SCT dan MCQ dilakukan dengan melihat *internal consistency* menggunakan koefisien *Alpha*. Hasil pengujian reliabilitas SCT mendapatkan koefisien sebesar 0.850. Sementara koefisien reliabilitas MCQ adalah 0,7.

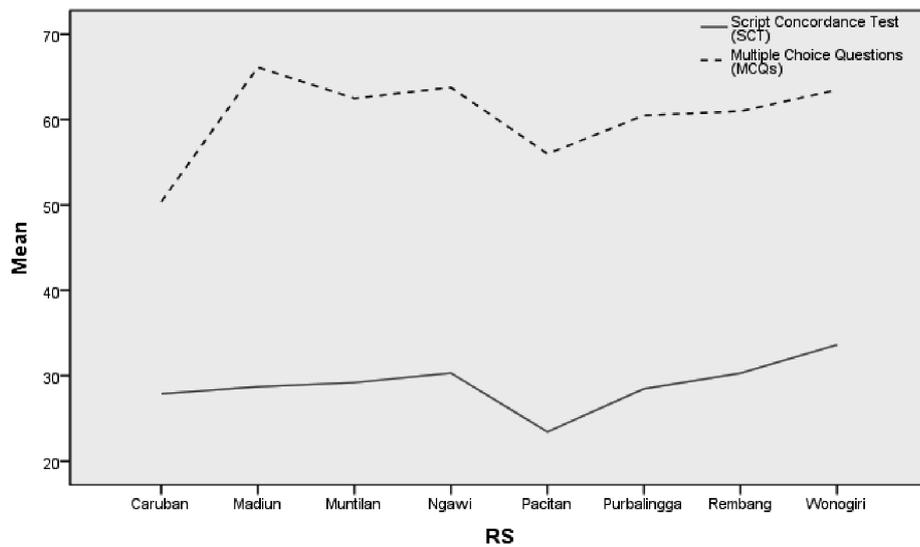
Analisis reliabilitas yang dilakukan pada set soal SCT dilakukan menggunakan konsistensi internal dengan *cronbach alpha*. Soal SCT menunjukkan hasil *cronbach alpha* = 0,850, lebih tinggi dibanding penelitian terdahulu (*cronbach alpha* = 0,75).<sup>13</sup> Hasil reliabilitas MCQ didapatkan nilai *cronbach alpha* = 0,75, sesuai dengan penelitian terdahulu (*cronbach alpha* antara 0,75 dan 0,95).<sup>13,23</sup>

### **Concurrent Validity**

Analisis *concurrent validity* diawali dengan pemberian set soal MCQ dan SCT pada mahasiswa di akhir stase Ilmu Kesehatan Anak. Sehari setelah mengerjakan MCQ, mahasiswa dihadirkan kembali untuk mengerjakan soal SCT.

Pengujian distribusi hasil pengukuran MCQ mendapatkan koefisien *Kolmogorov-Smirnov* = 1.918 dengan probabilitas ( $p$ ) = 0.001, perolehan  $p < 0.05$  menandakan data tidak berdistribusi normal, sehingga pengujian *concurrent validity* dilakukan dengan korelasi *rank spearman*.

Hasil pengujian mendapatkan koefisien korelasi *rank spearman* = 0.291 dengan probabilitas = 0.058, menandakan tidak signifikan pada toleransi kesalahan 5% namun signifikan pada toleransi kesalahan 10%. Hal ini berarti memiliki hubungan signifikan yang lemah. Berarti hasil pengukuran SCT memiliki konkordansi lemah dengan hasil pengukuran MCQ.



Gambar 2. Kesesuaian Hasil Pengukuran SCT dan MCQ

Gambar 2 menunjukkan kedua hasil pengukuran kurang memiliki kesejajaran yang sempurna, yang menandakan perbedaan karakter hasil pengukuran. Gambar juga menunjukkan hasil pengukuran MCQ lebih tinggi dibanding SCT.

Soal MCQ dan SCT selanjutnya dilihat *concurrent validity*-nya dengan memberikan kedua set soal kepada mahasiswa. Hasil dari *concurrent validity* ini dianalisis dengan uji *Spearman* karena data terdistribusi tidak normal dengan hasil bahwa MCQ dan SCT tidak mempunyai korelasi pada tingkat kesalahan 5%, akan tetapi berkorelasi lemah pada tingkat kesalahan 10%. Hal ini menunjukkan tidak terdapatnya *concurrent validity* yang kuat antara MCQ dan SCT.

Hal ini dimungkinkan karena perbedaan yang terdapat pada perancangan, proses berpikir dalam mengerjakan, dan proses skoring yang terjadi pada MCQ dan SCT. SCT dan MCQ merupakan alat ukur yang dapat digunakan untuk menilai kemampuan penalaran klinik. MCQ dapat mengukur kemampuan penalaran klinik yang terkait dengan aktivasi *high order clinical reasoning* dalam penyelesaian masalah (*problem solving*).<sup>20</sup> SCT merupakan alat ukur yang dapat menilai kemampuan penalaran klinik untuk keadaan *uncertainty*. SCT mempunyai kemampuan menilai penalaran klinis lebih jelas daripada MCQ.<sup>21</sup> Dalam hal ini SCT mempunyai kemampuan untuk aktivasi pada *illness script* proses dalam penalaran klinik dan proses *clinical data interpretation*.<sup>14,22</sup>

Dalam analisis kongruensi yang telah dilakukan terhadap MCQ dan SCT terdapat korelasi yang lemah antara SCT dan MCQ. Sesuai dengan teori *script*, bahwa pada *real setting*, para klinisi mengaplikasikan pengetahuan lama mereka untuk memahami munculnya gejala klinis pasien. Selanjutnya para klinisi akan menggunakan pengetahuan mereka untuk menolak atau menyetujui hipotesis mereka tentang suatu penyakit berdasarkan informasi keadaan klinis yang tampak.<sup>21</sup> Pendapat klinisi ini kemudian menjadi dasar skoring pada SCT. Salah satu kelemahan dan juga sekaligus kelebihan yang terdapat pada SCT adalah terkait proses skoring. Pakar akan menilai berdasarkan hal yang telah dialaminya. Hal ini akan menimbulkan perbedaan dalam interpretasi kemampuan mahasiswa. Akan tetapi di satu sisi, *clinical experience* ini akan menjadi hal yang baik untuk dijadikan bahan evaluasi bagi kemampuan mahasiswa dalam kemampuan mereka sesuai dengan keadaan yang sebenarnya (*real setting* dan *ill defined problems*)

Hasil konkordansi lemah (rendah) antara SCT dengan MCQ bukan menjelaskan salah satu lebih tidak akurat, melainkan lebih menjelaskan pada beda penekanan hal yang dinilai. Dalam hal ini SCT lebih fokus pada penalaran klinis yang terkait dengan *uncertainty* sedangkan MCQ pada *decision making* sebuah diagnosis atau terapi. Pada MCQ mahasiswa akan dinilai kemampuannya menegakkan

diagnosis, terapi, dan tatalaksana dari sesuatu yang pasti. Kepastian ini dikarenakan bahwa MCQ dalam penyusunannya harus didasarkan pada hal yang pasti baik terkait teori dan aplikasinya. Sehingga dalam penyelesaiannya pun MCQ membutuhkan kemampuan secara teori dan kognitif untuk dapat diaplikasikan dalam penyelesaian soal.

Pada SCT, proses mengerjakan didasarkan pada pembentukan *illness script* seorang dokter. *Illness script* akan terbentuk dari paparan kasus, kemampuan mendiagnosis, dan kemampuan memberikan terapi sesuai dengan keadaan pasien.<sup>19</sup> Dalam hal ini SCT akan menilai kemampuan mahasiswa dalam mengolah informasi sesuai dengan data yang tersedia. Data yang akan dilihat dan dievaluasi adalah terkait dengan *diagnostic knowledge*, *treatment knowledge*, dan *investigation knowledge*.

Kongruensi yang lemah antara SCT dan MCQ disebabkan juga karena beda perancangan soal antara MCQ dan SCT. Pada kasus MCQ *stem* soal dirancang untuk kasus yang sejalan dengan teori. Variasi yang bisa dibuat dengan diagnosis tertentu adalah sesuai dengan teori yang ada. Misalnya pada pembuatan kasus anemia yang sering terjadi pada anak adalah anemia defisiensi besi. Maka pada pembuatan kasusnya akan dibuat kasus anemia defisiensi besi pada anak yang berjumlah 1 soal. Sementara pada kasus SCT, mahasiswa akan diberikan ilustrasi kasus yang berisikan *ill defined problems* tentang anemia, dan kemudian diberikan *differential diagnosis* tentang anemia tersebut menjadi jumlah soalnya, misalnya terdapat 4 diagnosis yang mungkin, maka jumlah soal pada SCT menjadi 4 buah soal. Sehingga jumlah soal pada MCQ dan SCT tentang sebuah *focus (diagnostic, treatment, atau investigation knowledge)* tidaklah sama.

Selain itu, variasi dalam perancangan soal MCQ mempunyai lebih banyak kemungkinan soal. Hal ini dikarenakan ketika mahasiswa dinilai kemampuan menegakkan diagnosis, maka hal-hal pendukung lain bisa saja ditanyakan, misalnya patogenesis, masa inkubasi, faktor risiko, dan lain sebagainya. Sementara pada SCT harus berupa kasus *uncertain* pada pendekatan 3 aspek saja.

Pada saat mahasiswa mengerjakan SCT, untuk bisa memilih jawaban, mahasiswa harus mengetahui

dan mempunyai *clinical experience* untuk fokus diagnosis tersebut. Sehingga pada 1 *stem* soal, mereka harus mengetahui kelima diagnosis banding yang disediakan pada soal. Hal ini tidak terjadi pada MCQ, karena untuk memilih 1 diagnosis pasti, memungkinkan mahasiswa hanya menebak atau mengira dari pilihan yang dirasa benar.

### **Persepsi mahasiswa terhadap penggunaan SCT sebagai alat ukur penalaran klinik**

Berdasarkan kuesioner yang dibagikan kepada mahasiswa, menurut pendapat mahasiswa, 33 mahasiswa (76,74%) menyatakan setuju SCT dapat menilai penalaran klinis. Beberapa alasan yang dikemukakan bahwa SCT dapat mengarahkan diagnosis dan terapi, kasusnya dapat mencerminkan kemampuan petugas kesehatan dalam memberikan terapi.

Terkait dengan kemampuan SCT dalam menilai kemampuan interpretasi data klinik, 80,29 % mahasiswa menyatakan setuju, karena SCT membantu mahasiswa berpikir kritis, aplikatif dan mirip dengan kasus klinik yang dihadapi. Selain itu 19,21% (7 orang) mahasiswa menyatakan cukup. Beberapa alasan yang dikemukakan adalah bingung akan pilihan jawaban, ada perbedaan antara data klinik yang ditemukan dan teorinya.

SCT dapat menilai kemampuan mahasiswa menentukan terapi pada penyakit terkait kesehatan anak menurut 81,39% (35 orang) mahasiswa. Alasan yang mereka kemukakan antara lain adalah karena dapat menggambarkan kasus yang sering ditemui, informasi kasus dan kebutuhan terapi yang akan diberikan. Sebagian mahasiswa (18,60%) menyatakan SCT kurang tepat menilai keputusan memberikan terapi karena terapi di masing-masing RS dirasa berbeda protapnya. Untuk menentukan diagnosis, 34 mahasiswa (79,09 %) menyatakan setuju dengan metode SCT karena bentuk kasus mirip dengan yang terjadi di masyarakat.

### **KESIMPULAN**

*Concurrent validity* antara *Script Concordance Test* dengan *Multiple Choice Questions* tidak memiliki hubungan yang kuat pada penilaian penalaran klinik. *Script Concordance Test* tidak untuk menggantikan *Multiple*

Choice Questions melainkan untuk melengkapi strategi menilai penalaran klinis yang komprehensif dan mendekati pada *real setting*.

#### UCAPAN TERIMAKASIH

Penulis menghaturkan terimakasih kepada dr. Gandes Retno Rahayu, M.Med.Ed, Ph.D yang telah memberikan banyak bimbingan, dorongan motivasi dan masukan pada penelitian ini. Ucapan terimakasih juga penulis ucapkan kepada dr. Suryadi, SU, PA(K), MHPE yang telah memberikan masukan dalam penelitian ini. Dalam penelitian ini penulis sangat terbantu dengan peran dr. Tien Budi Febriani, M.Sc, Sp.A, yang telah menemani peneliti sebagai pakar Ilmu Kesehatan Anak.

#### PERSETUJUAN ETIKA

Penelitian ini telah mendapatkan persetujuan dari Komisi Etik dan Kesehatan FK UGM.

#### DAFTAR PUSTAKA

1. Higgs J, Jones M. *Clinical Reasoning in the medical profession*. Butterworth Heinemann: Elsevier; 2000.
2. Rahayu GR, McAleer. Clinical reasoning of Indonesian medical students as measured by diagnostic thinking inventory. *South East Asian Journal of Medical Education*, 2008;2(1):42-7.
3. Lee A, Joynt GM, Lee AKT, Groves M, Vlantis AC, Ronald CW, Colman SC, Cindy ST. Using illness script to teach clinical reasoning skill to medical students, *Fam Med*; 2010;42(4):255-61.
4. Vleuten C, Charlin B, Roy L, Brailovsky C, Goulet F. The script concordance test: a tool to assess the reflective clinician. *Teaching and Learning in Medicine*, 2000;12(4):189-95.
5. 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.
6. 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(4):e162-8.
7. Miller D, Linn RL, Gronlund NE. *Measurement and assessment in teaching*. Pearson education Internasional; 2009.
8. Elstein AS. Beyond multiple choice question and essay: the need for a new Gossway to assess clinical competence, *Academic Medicine*, 1993;68(4):244-9.
9. Sibert L, Charlin B, Corcos J, Gagnon R, Grisse P, Vleuten CVD. Stability of clinical reasoning assessment result with the script concordance test across two different linguistic, cultural, and learning environments. *Medical Teacher*, 2002;24(5):522-7.
10. Hernandez SZ, Rodriguez MC, Lopez VO, Montero RE, Cordoba AR. Validation of the script concordance test as an instrument to assess clinical reasoning of residents in pediatric emergency medicine in Mexico. *Bol Med Hosp Infant Mex*, 2011;68(5):369-73.
11. Duggan P. Development of script concordance test using an electronic voting system. *Ergo*, 2007;1:35-41.
12. Charlin B, Roy L, Brailovsky C, Goulet F. The Script concordance test: A tool to assess the reflective clinician. *Teaching and Learning in Medicine*, 2000;12(4):189-95.
13. Duggan P, Charlin B. Summative assessment of 5th year medical students' clinical reasoning by script concordance test: requirement and challenges. *BMC Medical Education*, 2012;12(29).
14. Lubarsky S, Charlin B, Cook DA, Chalk C. Script concordance testing: a review of published validity evidence. *Medical Education*, 2011;45:329-38.
15. Ayre C, Scally AJ. Critical values for lawshe content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling Development*, 2014;47(1):79-86.
16. Lambert D, Lines D. *Understanding Assessment, Purpose, Perceptions, Practice*. Routledge Falmer; 2000.
17. Cohen AS, Wollack JA. Helpful tips for creating reliable and valid classroom test: getting started—The test blueprint. *The Learning Link*, 2003;3(4):1-2.
18. Considine J, Botti M, Thomas S. Design, format, validity and reliability of multiple choice questions for use in nursing research and education. *Collegian*, 2005;12(1):19-24.
19. Lubarsky S, Dory VR, Duggan P, Gagnon R, Charlin B. Script concordance testing: from theory to practise: AMEE guide no 75. *Medical Teacher*, 2013;35:184-93.
20. Coderrre SP, Harasym P, Mandin H, Fick G. The impact of two multiple choice-question formats on the problem solving strategies used by novices and expert. *BMC Medical Education*, 2004;4:23.
21. Kelly WF, Durning S, Denton GD. Comparing a Script Concordance Examination to a Multiple-

- Choice Examination on a Core Internal Medicine Clerkship. *Teaching and Learning in Medicine*, 2008;24(3):187-93.
22. harlin B, Brailovsky C, Leduc C, Blouin D. The diagnosis script questionnaire: A new tool to assess a spesific dimension of clinical competence. *Advances in Health Science Education*, 1998;3:51-8.
23. Kania RE, Verrilaud B, Tran H, Gagnon R, Kazitani D, Huy PTB, Herman P, Charlin B. Online script concordance test for clinical reasoning assessment in otorhinolaryngology: the association between performance and clinical experience. *Arch Otorhinolaryngology Head Neck Surg*, 2011;137(8):751-5.