

UJI DIAGNOSTIK SOAL PILIHAN GANDA BERBASIS KASUS DAN *SCRIPT CONCORDANCE TEST* DALAM PENENTUAN KOMPETENSI PENALARAN KLINIS MAHASISWA D3 KEBIDANAN

Lia Nurwiliani

Dosen Kebidanan Politeknik Kesehatan TNI AU Ciumbuleuit Bandung

ABSTRAK

Kemampuan penalaran klinis sangat penting untuk mendukung kompetensi bidan dalam mengambil keputusan klinik. Situasi gawat darurat seringkali muncul tidak terduga, seperti halnya kasus-kasus perdarahan. Kompetensi penalaran klinis bidan dalam mengatasi perdarahan tersebut perlu dikuasai dan diuji dengan alat uji yang baik. Dua set soal tentang perdarahan dalam bentuk Soal Pilihan Ganda Berbasis Kasus (SPGBK) dan *Script Concordance Test* (SCT) telah dikembangkan. Perlu diketahui bagaimana nilai diagnostik SCT, SPGBK dalam menyaring mahasiswa yang belum kompeten dalam penalaran klinis. Penelitian potong silang dilakukan terhadap 62 subjek penelitian, yakni Mahasiswa tingkat 3 program Studi DIII kebidanan Poltekes TNI AU Ciumbuleuit yang telah lulus mata kuliah obstetri patologi dan ASKEB IV. Mereka yang sudah lulus dinilai sudah kompeten dalam penalaran klinis menurut hasil ujian dengan Soal Pilihan Ganda (SPG), termasuk materi ujian perdarahan. Responden diminta untuk mengerjakan 3 set soal, yakni SPG, SPGBK dan SCT. Nilai dari setiap jenis soal dianalisis untuk membandingkan kinerja tiap jenis soal dengan uji diagnostik. Korelasi antara nilai mahasiswa dalam dua ujian (SPGBK & SCT) diuji dengan uji korelasi *Pearson Product Momen*. Kesepakatan antara jenis soal dalam menilai penalaran klinis, diuji dengan metode *Bland-Altman*. Hasil hitung sensitifitas SPGBK 72,5%, spesifisitas 45%, Nilai Duga Positif (NDP) 70,7%, Nilai Duga Negatif (NDN) 47,6%, Rasio Kemungkinan Positif (RKP) 1,3 dan Rasio Kemungkinan Negatif (RKN) 0,6. Hasil hitung sensitifitas SCT 67,5%, spesifisitas 27,3%, NDP 62,8%, NDN 31,6%, RKP 0,93 dan RKN 1,19. Nilai r-hitung uji korelasi SPGBK dan SCT adalah $r=0,08$ dengan nilai $p=0,496$. Hasil uji *Bland Altman* menunjukkan selisih simpangan kesepakatan SPG dan SPGBK adalah 68,9, selisih simpangan kesepakatan SPG dan SCT adalah 66,9 dan selisih simpangan kesepakatan SPGBK dan SCT adalah 43,3. Batas simpangan yang dapat diterima adalah 44. Penelitian ini menunjukkan bahwa SPGBK dan SCT tidak dapat menggantikan SPG dalam menguji penalaran klinis. Tidak terdapat korelasi antara skor SCT dan SPGBK diduga karena karakteristik alat uji yang berbeda. Meskipun demikian SCT dan SPGBK sama-sama dapat digunakan dalam menguji penalaran klinis.

Kata kunci : Penalaran, Klinis, SPG, SPGBK, SCT

ABSTRACT

Clinical reasoning skills is crucial to support decision-making competence of midwives in the clinic. Especially emergency in situations. The most common emergency situation bleeding care. The competences must evaluate with appropriate assesment method. SPGBK and SCT has been recognized able to measure clinical reasoning competencies among health profession students. However there two methods diagnostik value have been explored yet. Cross sectional study conducted on 62 subjects, wich are students D3 level III Midwife Poltekes TNI AU Ciumbuleuit who have already pass bleeding corse which is a matter of the bleeding. Respondents were asked to do 3 sets of questions, wich are SPG, SPGBK and SCT. The result each of assessment methods are analyzed with diagnostic test approach and the Bland Altman method for suitability aspect. respondents is result analisis using a diagnostic test to observe sensitivity, spesivity, positive predictive value (NDP), negative predictive value (NDN) and negative likelihood ratios (RKN) and positive likelihood ratio (RKP). Validity and reliability test has been done before the result scoring. SPGBK is an equivalent to an SPG as clinical

reasoning assessment. The some result being occurred in SCT. There is also no correlation between SCT and SPGBK score. Otherwise there is agreement between SCT and SPGBK score due to clinical reasoning. The purpose of SPG is only measure factual knowledge aspect. It is clearly different with SCT and SPGBK. So because of this, there are no agreement between SCT and SPG, also between SPGBK and SPG. Due to the some ability of SCT and SPGBK to asses clinical reasoning, so there are agreement found between there two methods. Reseacher also found that is more difficult and need more time rather than conventional test. Other aspect that need to be concerned in this research is the quality of conventional test that been used in Poltekes TNI AU Ciumbuleuit, reseacher found there is no effort for item review.

Key word: *Clinical, Reasoning, SCT, SPGBK, Konvensional Test*

Pendahuluan

Sebab kematian ibu terbesar adalah perdarahan, infeksi, dan hipertensi (Departemen Kesehatan Republik Indonesia, 2007). Ketiga penyebab utama kematian ibu tersebut merupakan kasus emergensi yang membutuhkan penanganan secara cepat dan tepat, terutama kasus perdarahan. Ada dugaan bahwa keterlambatan penegakan diagnosis, pengambilan keputusan klinis dan pemecahan masalah pada kasus perdarahan memberikan andil sebesar 27% dari semua kematian ibu (Hernawati, 2011). Berdasarkan data RISKESDAS 2010 terdapat 82,2% persalinan ditolong oleh tenaga kesehatan dan sebanyak 62,1% diantaranya ditolong oleh bidan. Hal ini menunjukkan bidan berperan penting sebagai mitra perempuan dan tenaga kesehatan profesional strategis dalam peningkatan kesehatan ibu dan anak di Indonesia (Naskah Akademik Sistem Pendidikan Kebidanan, 2012).

Kompetensi paling utama yang harus dimiliki oleh tenaga kesehatan profesional adalah penalaran klinis (Dory et al, 2012). Beberapa bentuk penilaian penalaran klinis yang digunakan pada pendidikan tenaga kesehatan di antaranya adalah SPG, esai, *Script Concordance Test* (SCT) (Brugge et al, 2012). Kemampuan penalaran klinis menunjang keterampilan bidan dalam memecahkan masalah dan mengambil keputusan. Sebagai tenaga kesehatan profesional kemampuan penalaran klinis wajib dimiliki oleh setiap bidan. Lambatnya penurunan AKI dan AKB di Indonesia diduga terjadi sebagai akibat kompetensi kemampuan penalaran klinis bidan yang rendah dan berkontribusi pada “3 terlambat” (Depkes, 2011). Kondisi ini menunjukkan bahwa mungkin lulusan bidan yang sudah dinyatakan kompeten, berpraktek dengan ijin pemerintah

sesungguhnya belum benar-benar kompeten dalam pengambilan keputusan klinik. Timbul dugaan bahwa dalam uji kompetensi yang dilakukan oleh institusi pendidikan maupun organisasi profesi menggunakan metode evaluasi yang tidak tepat dalam aspek penalaran klinis. Pada penelitian ini peneliti ingin menguji dua alternatif alat uji kemampuan penalaran klinis, yakni SPGBK dan SCT.

Fokus SPGBK adalah untuk menilai keterampilan pengambilan keputusan dalam menegakan diagnosis dan penatalaksanaan.⁵ Begitu pula dengan *Script Concordance Test* (SCT) sebagai alternatif format instrumen untuk penilaian kompetensi peserta didik dalam dimensi penalaran klinis untuk pengambilan keputusan klinik, yang digunakan dalam praktik tatalaksana pasien sehari-hari (Fournier J., Demeester A., Charlin B, 2008).

SPGBK dan SCT merupakan bentuk soal yang berbasis kasus. Setiap kasus diikuti serangkaian pertanyaan. Untuk kasus dan pertanyaan pada SPGBK harus memberikan arahan yang jelas mengenai yang harus dilakukan peserta uji (Wood & Cole, 2004). Berbeda dengan SCT, kasus yang di gambarkan berada pada zone abu-abu. Untuk setiap kasus dalam zone tersebut dapat menimbulkan interpretasi berbeda, dengan pilihan solusi yang dapat dibenarkan menurut situasi dalam kasusnya. Para ahli sekalipun dapat mengambil satu pilihan berlainan, baik itu karena tidak semua data tersedia atau karena beberapa opsi dapat dibenarkan. SCT umumnya digunakan dalam menentukan diagnosis, pengobatan, isu-isu manajemen, dan dapat digunakan dalam aspek yang lebih khusus seperti masalah etika atau profesionalisme. Pilihan jawaban dalam SCT banyak dipengaruhi oleh wawasan internalnya.⁷

Gambar 1. Contoh SCT

Ny Ari, seorang G1P0A0, hamil 38 minggu mengeluh mules-mules sejak tadi malam dan bertambah kuat. Belum keluar air dari jalan lahir, demikian pula lendir maupun darah.

Pernyataan 1	Pernyataan2	Maka kekuatan hubungan antara informasi pada kolom 1 dan kolom 2 adalah :	
Jika anda menduga Ny Ari adalah parturien	Pada pemeriksaan luar didapatkan kontraksi yang berulang setiap 3-5 menit, intensitas lemah, selama 20-25 detik	- 2	Dapat disingkirkan
		- 1	Kecil kemungkinannya
		0	Tidak dapat dibuat kesimpulan
		+1	Besar kemungkinannya
		+2	Hampir dapat dipastikan
Jika pada pemeriksaan dalam didapatkan portio tebal, aksial, pembukaan 2 jari	Saya akan menyarankan Ny Ari untuk pulang	- 2	Sangat membahayakan
		-1	Membahayakan
		0	Tidak berpengaruh
		+1	Membantu
		+2	Sangat membantu

Ket: = kasus = hipotesis = pilihan jawaban

Tips penting dalam membuat SCT yaitu: Pertama, fokus pada hal pokok, yaitu pada data positif atau negatif yang berguna dalam menentukan solusi. Kedua sebarakan jawaban pada setiap pilihan dari skala Likert. Jangan sampai jawaban yang mungkin dari semua item pertanyaan adalah -1 dan +1 karena hal ini dapat dengan mudah di tebak oleh peserta ujian. Ketiga, pilihan pada skala likert harus didefinisikan secara jelas untuk mencegah adanya persepsi yang memiliki arti ganda. Keempat, arti dari pilihan nol pada skala, tidak memiliki dampak positif maupun negatif pada pilihan yang di pilih. Pilihan nol bukan pilihan yang aman di mana peserta test mengemukakan pendapat “saya tidak tahu” pada sebuah jajak pendapat yang menggunakan skala likert. Kelima, arti dari setiap pilihan pada skala likert berbeda-beda sesuai dengan jenis pertanyaan yang diajukan apakah

diagnostik, pemeriksaan, pengobatan atau manajemen.

Perbedaan selanjutnya antara SPGBK dan SCT adalah terletak pada cara pembobotan. Untuk SPGBK cara pembobotan adalah 1 jika jawaban benar dan 0 jika jawaban salah. Jawaban benar ditentukan oleh pembuat soal dimana hanya ada satu jawaban yang benar. Berbeda dengan SCT jawaban benar di tentukan melalui konsensus. Dimana pilihan yang terbaik ditentukan atas proporsi jawaban panel yang memilih satu jawaban tertentu. Anggota panel diminta untuk menyelesaikan SCT secara individual, dan jawaban mereka dijadikan kunci jawaban dari SCT dengan demikian skor SCT berkisar antara 0 s/d 1. Fournier J., Demeester A., Charlin B, 2008). ⁸Berikut contoh analisis pembobotan jawaban dari panel ahli beranggotakan 18 orang.

Tabel 1. Contoh Distribui dan Analisis Pembobotan Jawaban Panel Ahli

	Pilihan Jawaban				
	-2	-1	0	1	2
Distribusi jawaban panel pada satu pertanyaan	15	2	1	0	0
Skor (proporsi)	15/15	2/15	1/15	0	0
Nilai jawaban peserta ujian pada satu pertanyaan	1	0,13	0,06	0	0

Ket: Pada sebuah soal terdapat jawaban 18 anggota panel ahli, 15 orang memilih (-2), 2 orang memilih (-1) dan 1 orang memilih (0). Nilai untuk jawaban (-2) adalah 1, skor jawaban

dengan dasar perhitungam 15/15, untuk pilihan (-1) skor jawaban adalah 0,13 dengan perhitungan (2/15), sedangkan pilihan (0) skor jawaban adalah 0,06 dengan perhitungan (1/15). Untuk

jawaban yang tidak di pilih yaitu (+1) dan (+2) maka nilainya adalah 0. Dengan metode ini semua pertanyaan memiliki rentang yang sama yaitu 0 sampai dengan 1. Skor yang diperoleh dari setiap butir soal diakumulasikan untuk memperoleh skor total tes. Jumlah ini kemudian dibagi dengan jumlah pertanyaan dan dikalikan dengan 100 untuk mendapat nilai persentase.

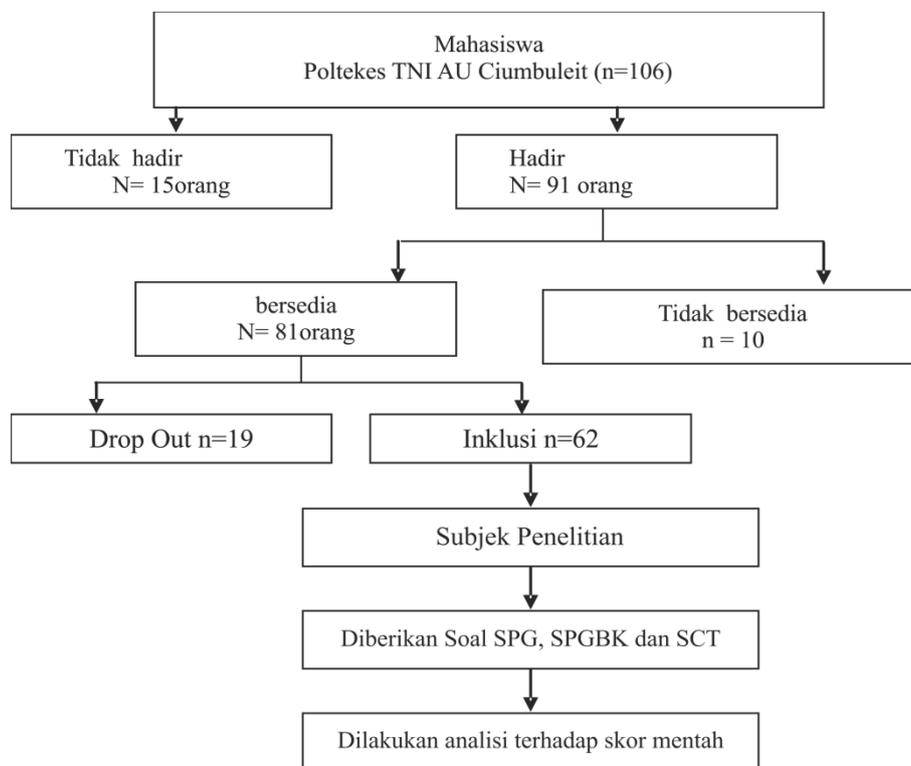
Dari uraian diatas dapat diketahui bahwa terdapat soal alternatif untuk menguji kemampuan penalaran klinis mahasiswa, namun perlu diketahui apakah soal alternatif tersebut menghasilkan proporsi kelulusan yang sama dengan SPG atau tidak.

Metode

Rancangan penelitian ini adalah penelitian potong silang. Pada tahap awal, dilakukan pemilihan SPG di bank soal poltekes TNI AU Ciumbuleuit, pembuatan SPGBK dan SCT. Tahap selanjutnya adalah melakukan uji validitas

isi SPGBK dan SCT oleh para pakar. Setelah di dapatkan pendapat para ahli dilakukan revisi terhadap soal, sesuai dengan masukan yang diberikan oleh para ahli. Setelah SPGBK dan SCT di revisi dilakukan pengambilan data terhadap mahasiswa Poltekes TNI AU. Pada hari pengambilan data dari 106 sampel, yang tidak hadir 15 orang, 10 orang tidak bersedia berpartisipasi sedangkan 19 orang hanya mengikuti 2 jenis test sehingga subjek penelitian hanya 62 orang.

Pada hari yang ditentukan 25 mahasiswa STiKes Bhakti Kencana hadir dan mengikuti ke 3 tes terebut, sebagai responden uji validitas dan reliabilitas soal. Mahasiswa Poltekes TNI AU Ciumbuleuit diwakili oleh hanya 62 orang yang mengikuti ke tiga tes tersebut secara lengkap. Sebanyak 19 orang mahasiswa lainnya hanya mengikuti satu atau dua jenis ujian, sehingga dikeluarkan dari subjek penelitian dan dianggap sebagai subjek tambahan untuk uji validitas dan reliabilitas dan dikeluarkan dari subjek penelitian.



Gambar 1. Subjek penelitian

Seharusnya semua soal diuji validitas dan reliabilitas sebelum diujikan kepada subjek penelitian. Namun mengingat keterbatasan

waktu, dan ijin pimpinan institusi STiKes Bhakti Kencana dan Poltekes TNI AU Ciumbuleuit maka semua soal langsung diujikan kepada

mahasiswa. Hanya soal yang dinyatakan valid, yang dianalisis dalam penelitian ini, cara pengambilan data adalah diberikannya ketiga bentuk soal dalam satu hari, dengan jarak waktu antara soal adalah 10 menit.

Skor mahasiswa di analisis dengan menggunakan uji diagnostik dengan batas kelulusan nilai mutu B pada penilaian acuan normatif.

Kemudian untuk uji korelasi menggunakan *pearson product moment* selanjutnya untuk uji kesepakatan menggunakan metode *bland altman*.

Hasil dan Pembahasan

Nilai mentah yang dicapai mahasiswa dalam setiap bentuk soal ditampilkan dalam tabel berikut:

Tabel 2. Pencapaian Penalaran Klinis

Jenis Soal	N	%	Proporsi Kelulusan Penalaran Klinis (batas lulus B)	Proporsi Tingkat keberhasilan Ujian (batas lulus D)	
SPG	A	4	6,5	35,5%	91,9 %
	B	18	29,0		
	C	23	37,1		
	D	12	19,4		
	E	5	8,0		
SPGBK	A	5	8,1	33,9%	88,7 %
	B	16	25,8		
	C	19	30,6		
	D	15	24,2		
	E	7	11,3		
SCT	A	5	8,1	30,7%	91,9%
	B	14	22,6		
	C	23	37,1		
	D	15	24,2		
	E	5	8,0		

Kompetensi penalaran klinis mahasiswa masih kurang karena tingkat kelulusan yang rendah. Pada penelitian ini didapati bahwa mahasiswa yang lulus dengan soal SPG, SPGBK maupun SCT bahkan kurang dari 50%, padahal peserta uji yang dijadikan sampel dalam penelitian ini merupakan mahasiswa yang sebelumnya telah dinyatakan lulus dengan SPG yang sama dengan SPG yang digunakan untuk menyatakan kelulusan sampel sebelumnya.

Banyak hal yang dapat menyebabkan rendahnya kelulusan siswa. Apabila ditinjau dari aspek penilaian diantaranya dari alat uji yang digunakan, maka satu set soal SPG dibuat tanpa mengacu pada *blueprint*. validitas konstruk SPG tidak diketahui padahal validitas konstruk merupakan kriteria kelayakan soal. Soal dengan validitas konstruk yang baik bermakna materi yang diajarkan dan merupakan tujuan pembelajaran (Wass V., Vleuten CVd., Shatzer J., Jones R,

2001).

Selain tidak dibuatnya *blueprint* tidak pernah pula dilakukannya analisis mutu soal. Sehingga sebelumnya tidak diketahui validitas ataupun reliabilitas soal SPG padahal validitas dan reliabilitas merupakan prinsip penilaian, yang menunjukkan bahwa suatu alat uji dapat mengukur apa yang harus diukur. Dari hasil uji validitas pun diketahui bahwa pada SPG dari 63 yang tersedia hanya 26 soal yang valid. Hal ini menunjukkan bahwa kualitas SPG yang tidak baik. Banyak hal yang dapat membuat SPG memiliki kualitas yang tidak baik, diantaranya kualitas Sumber Daya Manusia (SDM) dalam membuat soal. Berdasarkan hasil studi pendahuluan diketahui bahwa sampai dengan dibuatnya SPG yang digunakan dalam penelitian ini dosen yang membuat soal tidak pernah terpapar dengan pelatihan maupun standarisasi pembuatan soal.

Hal lain yang dapat menyebabkan tingkat kelulusan yang rendah, adalah pada saat

pengumpulan data pemberian soal diberikan dalam satu hari yang sama dengan jarak waktu 10 menit antara satu set soal dengan soal yang lain. Kemudian Metode penilaian yang baik adalah yang dapat memberikan pengaruh terhadap proses pembelajaran (Epstein, 2013). Akan tetapi dalam penelitian ini di temukan bahwa mahasiswa yang telah dianggap lulus dengan SPG

sebelumnya, gagal lulus dengan pengulangan soal yang sama. Hal ini menunjukkan dampak terhadap proses pembelajaran dari SPG belum terpenuhi. (Epstein, 2013).

Kemudian berdasarkan hasil uji diagnostik diketahui bahwa SPGBK dan SCT tidak dapat menggantikan SPG, seperti yang disajikan pada tabel berikut ini.

Tabel 3. Tabulasi Silang Uji Diagnostik SPGBK

SPGBK	SPG		Jumlah
	Tidak Lulus	Lulus	
Tidak Lulus	29	12	41
Lulus	11	10	21
Jumlah	40	22	

Ket : Sensitivitas = 72,5 %, Spesifisitas = 45%, NDP = 70,7%, NDN = 47,6%, RKP = 1,3, RKN = 0,6

Tabel 4. Tabulasi Silang Uji Diagnostik SCT

SCT	SPG		Jumlah
	Tidak Lulus	Lulus	
Tidak Lulus	27	16	43
Lulus	13	6	19
Jumlah	40	22	

Ket : Sensitivitas = 67,5 %, Spesifisitas = 27,3%, NDP = 62,8%, NDN = 20%, RKP = 0,93, RKN = 1,19

Pada tabel di atas diketahui bahwa hasil uji diaagnostik tidak baik karenatidak memenuhi syarat $RKP > 10$ dan $RKN < 0,1$. Banyak hal yang dapat menjadi penyebab diantaranya karena adanya perbedaan komposisi materi perdarahan yang diujikan antara SPG dan SPGBK atau SCT sehingga tidak dapat diperbandingkan sebagai *gold standard* untuk uji diagnostik. Kemudian pembuatan SPG tidak pernah dibuat *blueprint* terlebih dahulu menyebabkan materi soal yang di ujikan memiliki komposisi yang tidak teratur, terbukti dengan SPG yang hanya memiliki 3 soal yang dapat menguji kemampuan penalaran klinis. Sementara SPGBK dan SCT telah dirancang khusus untuk menguji kemampuan penalaran klinis yaitu pada tahap “*Knows How*”. Hal ini ditunjang pula dengan validitas isi SPG, diketahui bahwa SPG yang digunakan dalam penelitian ini hanya menguji (*recall*) daya ingat saja(Nurwiliani, 2013). Sehingga dapat kita ketahui bahwa SPG yang merupakan *gold standard* pada penelitian ini tidak memiliki kualitas yang baik untuk menguji penalaran klinis.

Hal lain yang dapat menyebabkan SPGBK tidak dapat menggantikan SPG adalah akibat perbedaan tingkat kesukaran dan daya pembeda mutu soal. Berdasarkan hasil analisis daya beda dan tingkat kesukaran soal di ketahui bahwa SPGBK lebih sulit dibandingkan SPG, akan tetapi proporsi SPGBK kurang dapat membedakan mahasiswa yang pintar dan tidak. Sebaliknya, meskipun berdasarkan analisis mutu soal SPG mempunyai daya pembeda lebih baik untuk membedakan mahasiswa yang pintar dan tidak, namun validitas isi yang dilakukan dengan cara survey yang dilakukan peneliti terhadap dosen bidan justru menunjukkan bahwa SPGBK yang digunakan dalam penelitian ini lebih unggul dan mampu menggambarkan kondisi yang sesuai dengan keadaan sebenarnya di klinik, mengukur kemampuan berpikir analitik, kemampuan pengambilan keputusan, serta membedakan mahasiswa yang pintar dan tidak pintar serta bersifat objektif (Nurwiliani I., Judistiani TD., Editors, 2013).

Berbeda dengan SCT yang tidak dapat dili-

hat daya beda dan tingkat kesukaran soalnya maka pada SCT dilakukan validitas isi yang dilakukan terhadap dokter spesialis Obstetri Ginekologi di Fakultas Kedokteran Universitas Padjadjaran, mereka berpendapat bahwa bentuk soal SCT “Baik sekali untuk menilai kemampuan penalaran klinis bagi klinisi karena selain wawasan keilmuan ada keterlibatan pengalaman klinis/ sarana prasarana pada fasilitas pelayanan kesehatan” (Judistiani, 2013). Sesuai dengan penelitian yang dilakukan oleh Amini dkk di Iran menyatakan bahwa bentuk soal SCT dapat mengukur kemampuan penalaran klinis (Amini et al, 2011)

Selain karena perbedaan mutu soal, nilai diagnostik SCT yang tidak baik dalam penelitian ini juga dapat disebabkan fakta bahwa format soal SCT adalah bentuk soal yang baru baik untuk para ahli yang jawabannya dijadikan kunci jawaban soal maupun untuk subjek penelitian.

Untuk menjawab hipotesis selanjutnya diketahui bahwa hasil uji korelasi menunjukkan nilai r hitung = 0,088 dengan nilai $p=0,496$. Dapat dinyatakan bahwa tidak ada korelasi antara nilai SCT dan nilai SPGBK. Dalam kondisi demikian mahasiswa yang nilai SPGBKnya besar belum tentu nilai SCTnya juga besar, begitu pula sebaliknya.

Hal ini terjadi, karena meskipun SPGBK dan SCT sama-sama merupakan alat uji penalaran klinis, akan tetapi terdapat perbedaan sifat soal, dimana SPGBK mengukur penalaran klinis pada kasus yang bersifat jelas, sedangkan SCT mengukur penalaran klinis pada kasus yang dapat menimbulkan interpretasi berbeda. Selain sifat kasus yang berbeda terdapat pula perbedaan pada cara penilaian, dimana pada SPGBK hanya ada satu jawaban yang benar dengan skoring 1 jika benar dan 0 jika salah. Sementara itu, SCT memiliki jawaban benar lebih dari satu dengan skoring berkisar 0 sampai dengan 1.

Uji *Bland Altman* dalam penelitian ini dilakukan untuk mengetahui kesepakatan antara SPGBK dan SCT, didapatkan bahwa SPGBK dan SCT memiliki kesepakatan, sehingga dapat saling menggantikan. Hal ini sesuai dengan teori Miller yang menyatakan bahwa SPGBK dan SCT dapat menilai kemampuan kognitif pada tingkat “*Knows How*”.¹¹ Pada tingkat “*Knows How*” kemampuan yang dinilai adalah kemampuan

pengambilan keputusan dan kemampuan pemecahan masalah, yang merupakan aplikasi dari penalaran klinis.

Setiap alat uji tentu saja memiliki kekurangan dan kelebihan, Tambunan berpendapat bahwa tidak ada penilaian tunggal yang dapat digunakan untuk menilai kompetensi (Tambunan & Soetjiningaih, 2011). Penggunaan lebih dari satu metode penilaian dapat mengkompensasi metode yang lain (Epstein, 2007).

Maka dari itu penting dibuatnya *blueprint*, agar tujuan pembelajaran yang ingin dicapai benar-benar dinilai, sehingga peserta uji yang lulus merupakan peserta uji yang benar-benar kompeten (Wass V., Vleuten CV., Shatzer J., Jones R, 2001). Berdasarkan studi pendahuluan yang dilakukan diketahui bahwa hingga SPG ini dibuat dosen-dosen yang membuat soal pun belum terpapar mengenai *blueprint*.

Kemampuan penalaran klinis merupakan kompetensi yang harus dicapai oleh seorang bidan. Kompetensi merupakan kemampuan seseorang pada situasi kehidupan nyata. Sehingga sesungguhnya kompetensi yang sebenarnya hanya dapat dinilai pada dunia kerja (Eipstein, 2007).

Berdasarkan survey yang dilakukan peneliti, SCT dianggap lebih unggul di bandingkan SPGBK, dalam hal kemampuan soal untuk menilai kemampuan pengambilan keputusan dan kemampuan berfikir analitik soal. Meskipun SCT dianggap lebih unggul, akan tetapi SCT juga dianggap lebih cocok diberikan pada mahasiswa tahap akhir, sedangkan SPGBK dianggap lebih fleksibel dan dapat diberikan pada mahasiswa tahap awal. Oleh karena itu dapat disimpulkan bahwa SCT dan SPGBK dapat saling menggantikan, akan tetapi akan lebih baik jika keduanya digunakan untuk saling melengkapi.

Kesimpulan

Penelitian ini menunjukkan bahwa SPGBK dan SCT tidak dapat menggantikan SPG dalam menguji penalaran klinis. Tidak terdapat korelasi antara skor SCT dan SPGBK diduga karena karakteristik alat uji yang berbeda. Meskipun demikian SCT dan SPGBK sama-sama dapat digunakan dalam menguji penalaran klinis.

Daftar Pustaka

- Amini M, Moghadami M, Kojuri J, Abbasi H, Ali Arhami Doolat Abadi, Molaee NA, et al. (2011). An innovative method to assess clinical reasoning skills: Clinical reasoning tests in the second national medical science Olympiad in Iran. *BMC Research Notes*. 2011;4(418).
- Bruggen Lv, Woudenbergh MM-v, Spierenburg E, Vos J. (2012). Preferred question types for computer-based assessment of clinical reasoning: a literature study. *Perspectives on Medical Education*;1(4).
- Departemen Kesehatan Republik Indonesia. (2007). Jakarta: profil kesehatan Indonesia 2007.
- Departemen Kesehatan Republik Indonesia. (2011). Refleksi Hari Ibu: Skenario Percepatan Penurunan Angka Kematian Ibu". Jakarta.
- Dory V, Gagnon R, Dominique, Vanpee, Charlin B. (2012) How to construct and implement script concordance tests : insights from a systematic review. *Medical Éducation* 16 November 2012;46:552-63.
- Epstein R. Assessment in Medical Education. (2007). *The New England Journal of Medicine*. 2007;356:387 - 96.
- Fournier J, Demeester A, Charlin B. (2008). Script Concordance Tests: Guidelines for Construction. *BMC Medical Informatics and Decision Making*. 8(1):18.
- Hernawati I. (2011). Analisis kematian Ibu di Indonesia tahun 2010. In: RI D, editor. Jakarta
- Judistiani TD. (2013) .Interview. Tasya M, Husein F, Tita, Vita, Susanto. H, Brathakusuma DS.
- Naskah Akademik Sistem Pendidikan Kebidanan di Indonesia. (2012). In: IBI, AIPKIND, editors. Jakarta: HPEQ,
- Nurwiliani L, Judistiani TD, editors. (2013). Persepsi Dosen Bidan tentang Bentuk Soal Pilihan Ganda Berbasis Kasus dan Script Concordance Test (SCT) sebagai Alat Uji Kasus Perdarahan dalam Pendidikan Diploma III Kebidanan. *Pertemuan Ilmiah Bidan.* Bandung; Bandung.
- Nurwiliani L. (2013). Interview. Judistiani TD, Susanti Y, Susanto H.
- Tambunan T, Soetjningaih, Supriyatno B. (2011). Sistem Evaluasi pada Pendidikan Dokter Berbasis Kompetensi.
- Wass V, Vleuten CVd, Shatzer J, Jones R. (2001). Assessment of clinical competence. *The Lancet* March 24, 2001;Vol 357.
- Wood T, Cole G. (2004). Developing Multiple Choice Questions for the RCPSC Certification Examinations.