PENGARUH KARAKTERISTIK PENGUJI TERHADAP DERAJAT KESESUAIAN ANTAR-PENGUJI *OBJECTIVE STRUCTURED CLINICAL EXAMINATION* (OSCE) DI PROGRAM STUDI D3 KEBIDANAN FAKULTAS KEDOKTERAN UNPAD

Khalidatunnur Andrianie, Rani Nurparidah, Yuni Pratiwi Susanti, Tina Dewi Judistiani

ABSTRAK

Reliabilitas ujian *Objective Structured Clinical Examination* (OSCE) dapat ditingkatkan dengan menilai derajat kesesuaian antar-penguji dalam memberikan penilaian ujian.

Tujuan penelitian ini adalah untuk menganalisis derajat kesesuaian antar-penguji OSCE dan menganalisis pengaruh karakteristik penguji terhadap derajat kesesuaian antar penguji OSCE.

Penelitian ini menggunakan rancangan penelitian potong silang. Sampel penelitian ini adalah seluruh dosen tetap bidan berjumlah 16 orang, 10 mahasiswa Program Studi D3 Kebidanan FK Unpad semester 6 dan 2 orang pasien simulasi. Analisis penelitian untuk melihat derajat kesesuain antar-penguji menggunakan Fleiss Kappa sedangkan untuk melihat pengaruh karakteristik penguji terhadap derajat kesesuaian antar-penguji digunakan uji t tidak berpasangan, uji Mann-Whitney dan uji Kruskal-Wallis.

Hasil penelitian didapatkan bahwa terdapat perbedaan derajat kesesuaian antar-penguji OSCE (nilai kappa <0,4), karakteristik penguji yaitu pengalaman klinik, pengalaman menjadi pengajar keterampilan yang diujikan, keikutsertaan dalam standardisasi OSCE dan pengalaman menguji OSCE berpengaruh terhadap derajat kesesuaian antar-penguji (p <0,005), sedangkan masa kerja penguji dan pengalaman dalam merancang pojok uji yang dinilai tidak berpengaruh terhadap derajat kesesuaian antar-penguji OSCE.

Kesimpulan penelitian ini adalah derajat kesesuaian antar-penguji di Program Studi D3 Kebidanan FK Unpad kurang baik dan karakteristik penguji yang berpengaruh terhadap derajat kesesuaian antar-penguji adalah pengalaman klinik, pengalaman menjadi pengajar keterampilan yang diujikan, keikutsertaan dalam standardisasi OSCE dan pengalaman menguji OSCE.

Kata Kunci : Derajat kesesuaian antar-penguji OSCE dan karakteristik penguji

ABSTRACT

Inter-rater reliability can be improved the reliability test of the Objective Structured Clinical Examination (OSCE).

The purpose of this study was to analyze OSCE inter-rater reliability and the influence of the examiner characteristics.

This design was a cross sectional. The samples were 16 midwife lecturer, 10 students of the 6 semesters Midwifery diploma program and 2 simulated patients. The data were analysed using Fleiss Kappa to obtain inter-rater reliability and unpaired t test, the Mann-Whitney, the Kruskal-Wallis for the influence of the examiner characteristics to inter-rater reliability.

The results showed that OSCE inter-rater reliability was poor (kappa <0.4). Examiner characteristics such as, clinical experience, clinical skills trainer experience, standardization involvement OSCE and examiner experience were influence inter-rater reliability (p <0.005), while the rater tenure and experience in station construction did not influence on inter-rater reliability.

The conclusion of this study was OSCE inter-rater reliability of the Midwifery Diplom Program of the Medical Faculty Padjadjaran University was poor and examiner characteristics that influence inter-rater reliability were clinical experience, clinical skills trainer experience, standardization involvement and OSCE examiner experience.

Keyword : OSCE inter-rater reliability, examiner characteristics

---

Latar Belakang

Penilaian hasil belajar merupakan bagian yang terpenting dalam kurikulum karena memberikan efek konstruktif yang kuat terhadap pembelajaran dan kurikulum. Penilaian hasil belajar dapat memberikan umpan balik kepada institusi pendidikan untuk meningkatkan mutu pendidikan dan kepada mahasiswa sebagai motivator mahasiswa untuk belajar.[1]

Tujuan penilaian hasil belajar menurut Amin dan Eng adalah menentukan apakah tujuan pembelajaran yang telah ditentukan sebelumnya telah tercapai, mendorong mahasiswa belajar, sertifikasi dan menilai kompetensi, mengembangkan dan evaluasi program pendidikan, mengerti proses pendidikan, dan memprediksi performa mahasiswa nantinya.[2]

Karakteristik penilaian hasil belajar terdiri dari: valid, artinya suatu metode penilaian dapat mengukur pencapaian hasil belajar yang seharusnya diukur; reliabel, artinya suatu metode penilaian hasil belajar secara konsisten membuahkan hasil penilaian yang sama bila digunakan untuk menilai mahasiswa oleh penilai dan pada tempat yang berbeda; *feasible;* dan mempunyai dampak pada pendidikan, artinya penilaian hasil belajar dapat memberikan umpan balik terhadap proses pembelajaran.[1, 3]

Penilaian hasil belajar dapat dilakukan terhadap aspek pengetahuan, kompetensi dan kinerja mahasiswa. Cara penilaian hasil belajar ini perlu disesuaikan dengan aspek yang ingin dinilai. *Multiple Choice Question* (MCQ), *essay*, *oral examination* digunakan untuk menilai aspek pengetahuan. *Objective Structured Clinical Examination (OSCE), simulated patient based test, Mini CEX, 360^0 evaluation* untuk kompetensi keterampilan klinik.

Berdasarkan piramid Miller, kompetensi seseorang terbagi dalam beberapa tahap, mulai dari tingkat tahu (*knows)* hingga tingkat terampil (*does)*. Keterampilan klinik menempati tingkat ketiga yaitu mampu mendemonstrasikan (*shows how.)*[1, 4]

OSCE diperkenalkan pertama kali oleh Harden dkk pada tahun 1975, ujian ini didesain untuk menilai hasil belajar mahasiswa pada keterampilan klinik dan kompetensi mahasiswa kedokteran tingkat akhir. Pengaturan OSCE menurut Harden, dkk terdiri dari 20-30 pojok uji (*station*) dengan alokasi waktu ujian antara 5-10 menit pada setiap pojok uji. Harden berpendapat bahwa ujian OSCE lebih objektif dan reliabel dalam menilai keterampilan mahasiswa dibandingkan dengan ujian konvensional dan berperan penting pada pengembangan keterampilan klinis yang efektif.[4-7]

Tujuan OSCE menilai kompetensi dan keterampilan klinis mahasiswa secara objektif dan terstruktur. Objektif maksudnya adalah setiap mahasiswa yang diuji dinilai dengan alat uji berupa daftar tilik yang sama, dengan kriteria kinerja yang terukur. Terstruktur maksudnya adalah bahwa sekumpulan mahasiswa diuji

dengan jenis tugas yang sama, dalam alokasi waktu ujian yang sama.[8-9]

Pada soal OSCE yang baik, validitas dan reliabilitas dapat terjaga apabila soal ujian mengacu pada kisi-kisi (*blueprint)* kompetensi yang akan diujikan. Aspek reliabilitas ujian dapat diperbaiki dengan menambah jumlah pojok uji atau menambah jumlah penguji dalam setiap pojok uji untuk mengurangi bias. Daftar tilik membantu tercapainya konsistensi dan stabilitas penilaian penguji, sebagai komponen penting untuk menentukan keakuratan ujian keterampilan.[1]

Terlepas dari pernyataan Harden, masih terdapat beberapa masalah dalam OSCE yang dapat mempengaruhi hasil ujian, yakni: faktor penguji, penggunaan pasien simulasi (*standardized patient*), bahan yang diujikan, sistem penilaian serta waktu pengayaan yang diberikan kepada mahasiswa. Objektivitas penguji OSCE dipengaruhi oleh pengalaman menguji sebelumnya, keikutsertaan dalam standardisasi penguji OSCE, pelatihan-pelatihan keterampilan yang pernah diikuti, pengalaman klinis penguji, pengalaman mengajar di bidang keterampilan yang dijadikan kompetensi ujian OSCE serta keterlibatan penguji dalam membangun OSCE. Sehingga agar menghasilkan penguji OSCE yang objektif maka syarat menjadi penguji OSCE adalah latar belakang pendidikan minimal S2, sudah berpengalaman menjadi instruktur keterampilan klinik baik di laboratorium maupun di klinik selama 1 tahun, telah mengikuti pelatihan yang terstandar sebagai penguji OSCE.[10] Penelitian OSCE yang dilakukan oleh *Medical Council of Canada* pada tahun 1992, menyatakan bahwa beberapa determinan penguji mempunyai pengaruh yang signifikan dalam 40% pojok uji yang dilaksanakan.[11-13]

Penelitian tentang pelaksanaan OSCE yang dilakukan pada 50 mahasiswa Fakultas Kedokteran Universitas Udayana menunjukkan 82% responden berpendapat bahwa penilaian OSCE masih subjektif dan hanya 18% responden mengatakan sudah sangat objektif. Banyak yang mengatakan bahwa penilaian OSCE ini masih dipengaruhi oleh perbedaan cara menilai dosen akibat pengaruh berbedanya karakter.[12]

Wilkinson menyatakan bahwa reliabilitas penguji perlu dievaluasi. Reliabilitas atau kesamaan persepsi antar-penguji dalam menilai dinyatakan sebagai derajat kesesuaian antar-penguji (*inter-rater reliability)*. Derajat kesesuaian antar-penguji yang baik menghindarkan persepsi yang berbeda atas keterampilan mahasiswa, sekalipun diuji oleh dua atau lebih penguji dalam menilai suatu kegiatan yang sama pada waktu sama. Jika nilai semua penguji sama, maka penilaian tersebut dikatakan objektif. Keterlibatan penguji dalam membangun pojok uji OSCE mempunyai hubungan yang signifikan terhadap derajat kesesuaian antar-penguji dalam menilai keterampilan klinis ($r = 0.23$, $p = 0.005$) tetapi pengalaman menguji OSCE dan pengalaman klinis penguji tidak mempunyai hubungan dengan derajat kesesuaian antar-penguji dalam menilai keterampilan klinis. Pengalaman menguji dan pengalaman klinis penguji akan mempunyai hubungan yang signifikan terhadap derajat kesesuaian antar-penguji ketika penilaian ujian menggunakan daftar tilik berskala karena pengalaman penguji akan berguna dalam menilai keseluruhan performa mahasiswa dalam melakukan keterampilan klinis.[13]

Hal lain yang dikaji dalam penelitian tersebut adalah korelasi lembar penilaian terhadap objektivitas ujian relatif kecil dibandingkan dengan faktor penguji, namun meningkatnya jumlah butir daftar tilik pada lembar penilaian dapat menurunkan validitas dan reliabilitasnya.[13]

Penelitian Hodges di Universitas Toronto menyatakan bahwa pengalaman klinis seorang penguji dapat mempengaruhi reliabilitas alat uji. Sebaliknya penggunaan daftar tilik tidak meningkatkan reliabilitas seorang penguji yang kaya pengalaman klinis. Reliabilitas seorang klinisi sebagai penguji lebih baik jika alat uji yang digunakan berupa sistem penilaian global

(*global scoring system)*. Hal ini terjadi karena dokter yang berpengalaman sering menilai pelaksanaan tugas klinis dari sisi efisiensi bukan semata-mata mencakup semua aspek yang tertera daftar tilik.[13-14]

Belum seluruh institusi pendidikan kebidanan di Indonesia menyelenggarakan ujian keterampilan klinik menggunakan metode OSCE karena belum semua institusi pendidikan mengikuti pelatihan ujian OSCE, sumber daya manusia yang belum memenuhi syarat sebagai penguji OSCE dan biaya penyelenggaraan ujian yang mahal.

Program Studi D3 Kebidanan FK Unpad telah berdiri sejak 2007, sampai saat dosen bidan yang ada berjumlah 17 orang, dan telah mengadakan ujian akhir program dengan metode OSCE selama lima kali, namun selalu lebih dari 70 % mahasiswa perlu menjalani remedial. Pada kenyataannya proses pembelajaran di Program Studi D3 Kebidanan sudah mengikuti metode *Student Centered Learning* (SCL), namun ketidaklulusan pada ujian OSCE kemungkinan dipengaruhi oleh faktor penguji, faktor mahasiswa dan faktor alat uji.

Penguji OSCE yang ada di Program Studi D3 Kebidanan FK Unpad sangat bervariasi baik dari usia, lama kerja, pengalaman klinik maupun pengalaman mengikuti pelatihan-pelatihan keterampilan klinik, pelatihan keterampilan melatih selain hal tersebut, para penguji belum pernah mendapatkan pelatihan sebagai penguji OSCE secara khusus sehingga syarat menjadi penguji OSCE di Program Studi D3 Kebidanan Unpad hanya telah mengikuti sosialisasi ujian OSCE dan mengajarkan keterampilan yang diujikan dalam ujian OSCE. Ketidaksesuaian antar-penguji dalam menilai sangat mempengaruhi hasil evaluasi keterampilan klinik sehingga perlu dilakukan evaluasi sejauh mana derajat kesesuaian antar-penguji pada ujian OSCE yang telah dilakukan.

Oleh karena itu, tujuan penelitian ini adalah untuk menganalisis derajat kesesuaian antar-penguji OSCE dalam menilai mahasiswa di Program Studi D3 Kebidanan FK Unpad dan menganalisis pengaruh karakteristik penguji yaitu masa kerja, pengalaman klinik penguji, pengalaman mengajarkan keterampilan yang diujikan dalan OSCE, pengalaman merancang pojok uji yang dinilai, keikutsertaan penguji dalam standardisasi OSCE dan pengalaman menguji OSCE terhadap derajat kesesuaian antar-penguji OSCE di Program Studi D3 Kebidanan Unpad.

Metode

Waktu penelitian pada bulan September-Desember 2013 bertempat di Program Studi D3 Kebidanan FK Unpad, sampel penelitian adalah seluruh dosen tetap bidan yang menjadi penguji OSCE (*total sampling*) berjumlah 16 orang, 10 orang mahasiswa Program Studi D4 Kebidanan FK Unpad yang berasal dari Program Studi D3 Kebidanan FK Unpad yang bersedia mengikuti simulasi ujian OSCE dan 2 orang pasien simulasi yang bersedia untuk menjadi pasien simulasi dalam simulasi ujian OSCE dengan desain penelitian potong silang analitik.

Cara kerja dan teknik pengumpulan data pertama-tama peneliti menentukan subjek penelitian berdasarkan kriteria inklusi, lalu meminta kesediaan kepada seluruh mahasiswa semester 6 untuk mengikuti penelitian ini, tetapi hanya 10 orang mahasiswa yang dipilih untuk pengikuti simulasi ujian OSCE. Selanjutnya peneliti meminta kesediaan pasien simulasi untuk melakukan simulasi ujian OSCE serta meminta seluruh penguji OSCE untuk menilai kompetensi mahasiswa pada simulasi ujian OSCE.

Peneliti membuat dokumentasi simulasi ujian OSCE pada seluruh area kompetensi yang harus dikuasai oleh bidan, yaitu asuhan kehamilan, persalinan, nifas, bayi baru lahir, bayi dan balita, pelayanan KB dan kegawatdaruratan maternal dan neonatal, yang terdiri dari 12 pojok uji. Simulasi Ujian OSCE ini dilakukan oleh 10 mahasiswa, di mana pada setiap pojok uji mahasiswa mengikuti ujian selama 10 menit dan dilakukan rekaman

menggunakan *video recorder*. Adapun kompetensi bidan yang diujikan adalah:

1) Pojok uji 1 : Asuhan kehamilan normal (pemeriksaan fisik, diagnosis)
2) Pojok uji 2 : Asuhan kehamilan patologi (konseling)
3) Pojok uji 3 : Asuhan persalinan (penatalaksanaan kala 2 dan 3)
4) Pojok uji 4 : Asuhan persalinan (penatalaksanaan penjahitan perineum)
5) Pojok uji 5 : Asuhan nifas normal (perawatan payudara)
6) Pojok uji 6 : Asuhan nifas patologi (pemeriksaan fisik dan diagnosis)
7) Pojok uji 7 : Asuhan bayi baru lahir (pemeriksaan fisik)
8) Pojok uji 8 : Asuhan neonatus, bayi dan balita (pemberian immunisasi)
9) Pojok uji 9 : Pelayanan KB (konseling alat kontrasepsi)
10) Pojok uji 10 : Pelayanan KB (pemasangan IUD Copper T)
11) Pojok uji 11 : Kegawatdaruratan maternal (penanganan atonia uteri)
12) Pojok uji 12 : Kegawatdaruratan neonatal (resusitasi BBL)

Pada pojok uji 2, 5, 8 dan 9 digunakan pasien simulasi, sedangkan pada pojok uji lainnya digunakan *phantoom*.

Diberikan kuesioner kepada penguji OSCE tentang karakteristik penguji lalu penguji OSCE akan menilai kompetensi mahasiswa dalam melakukan keterampilan klinik kebidanan dengan melihat rekaman video. Pada setiap pojok uji mahasiswa akan dinilai oleh 4 orang penguji yang menjadi sampel penelitian yang dipilih secara acak yaitu 2 orang penguji yang mempunyai masa kerja > 7 tahun dan 2 orang penguji yang mempunyai masa kerja < 7 tahun dengan menggunakan daftar tilik penilaian global, penguji yang pempunyai masa kerja > 7 tahun menjadi penguji 1 dan 2, penguji yang mempunyai masa kerja < 7 tahun menjadi penguji 3 dan 4. Penguji yang mempunyai masa kerja > 7 tahun mendapatkan kesempatan menilai 2-3 pojok uji, sedangkan penguji yang mempunyai masa kerja < 7 tahun mendapatkan kesempatan menilai 3-4 pojok uji. Pada saat menilai ujian, para penguji tidak boleh ada kontak satu sama lain agar penilaian yang dihasilkan murni berasal dari persepsi penguji tersebut bukan karena masukan dari penguji lainnya. Selanjutnya hasil penilaian kinerja mahasiswa dari masing-masing penguji pada setiap pojok uji yaitu sebanyak 480 lembar daftar tilik dan kuesioner karakteristik penguji akan dianalisis.

Pengujian hipotesis dilakukan untuk mengetahui derajat kesesuaian antar-penguji dengan menggunakan koefisien Fleiss Kappa dengan interpretasi nilai Kappa menurut Fleiss (1981) yang dikutip oleh von Eye A. dan Mun EY. adalah $k < 0,40$: buruk, $0,40 < k > 0,75$: baik, $k > 0,75$: sangat baik.[32]

Untuk melihat faktor karakteristik yang berpengaruh terhadap derajat kesesuaian antar-penguji pada 2 kelompok data digunakan uji t tidak berpasangan pada data terdistribusi normal dan Mann-Whitney pada data tidak terdistribusi normal, untuk melihat faktor karakteristik dominan yang berpengaruh terhadap derajat kesesuaian pada 3 atau lebih kelompok menggunakan Kruskal Wallis.

Hasil Penelitian

Terdapat perbedaan derajat kesesuaian antar-penguji OSCE di Program Studi D3 Kebidanan FK Unpad (tabel 1) didapatkan nilai kappa < 0.4 dan karakteristik yang berpengaruh teehadap derajat kesesuaian antar penguji adalah pengalaman klinik, pengalaman menjadi pengajar keterampilan yang diujikan, keikutsertaan dalam standardisasi OSCE, dan pengalaman menguji OSCE (tabel 2).

Tabel 1 Derajat Kesesuaian Antar-penguji OSCE dalam Menilai Keterampilan Klinik Mahasiswa pada Ujian OSCE

Pojok Uji	Fleiss Kappa	SE	IK 95%
Derajat Keseuaian Antar-penguji berdasarkan Hasil Nilai Ujian			
1	-0,06	0,05	-0,16 – 0,04
2	0,05	0,05	-0,05 – 0,15
3	-0,06	0,04	-0,13 – 0,03
4	0,04	0,05	-0,06 – 0,14
5	0,04	0,05	-0,06 – 0,13
6	-0,04	0,05	-0,14 – 0,01
7	-0,07	0,04	-0,14 – 0,01
8	0,09	0,06	-0,02 – 0,2
9	-0,07	0,05	-0,16 – 0,03
10	-0,05	0,03	-0,12 – 0,02
11	-0,06	0,01	-0,17 – 0,05
12	-0,04	0,04	-0,11 – 0,03
Derajat Keseuaian Antar-penguji berdasarkan Hasil Kelulusan			
1	-0,18	0,13	-0,43 – 0,07
2	-0,08	0,13	-0,33 – 0,17
3	-0,15	0,13	-0,40 – 0,11
4	0,11	0,13	-0,14 – 0,36
5	0,13	0,13	-0,12 – 0,39
6	-0,15	0,13	-0,4 – 0,11
7	0,13	0,13	-0,13 – 0,38
8	0,26	0,13	-0,01 – 0,51
9	0,05	0,13	-0,2 – 0,3
10	0,28	0,13	0,03 – 0,54
11	-0,08	0,13	-0,33 – 0,17
12	-0,37	0,13	0,11 – 0,62

Tabel 2 Pengaruh Karakteristik Penguji terhadap Hasil Penilaian Ujian OSCE pada Pojok Uji 1-12

Karakteristik Penguji	n	Rerata (s.b.)	Perbedaan Rerata (IK95%)	*P**
Masa kerja penguji				
< 7 tahun	10	75,5 (4,1)	0,5 (-3,2 – 4,2)	0,78
> 7 tahun	10	76,1 (3,7)		
Pengalaman menjadi pengajar keterampilan yang diujikan dalam OSCE				
Ya	10	76,8 (3,4)	6,7 (3,4 - 10)	0,001
Tidak	10	70,1 (3,7)		
Pengalaman dalam merancang pojok uji yang dinilai sekarang				
Ya	10	78,7 (2)	0,6 (-1,9 – 3,1)	0,62
Tidak	10	78,1 (3,2)		
Pengalaman menguji OSCE				
< 4 kali menguji	10	61,5 (3,8)	14,1 (10,6 – 17,7)	< 0,001
> 4 kali menguji	10	75,6 (3,8)		

Karakteristik Penguji	n	Median (minimum – maksimum)	*P*
Pengalaman klinik profesi			
< 3 tahun	10	76,5 (74 – 83)	*< 0,001***
> 3 tahun	10	58 (51 – 64)	
Keikutsertaan dalam standardisasi OSCE			
Jarang	10	63,5 (55 – 69)	< 0,001***
Sering	10	72,5 (67 – 77)	
Selalu	10	51,5 (47 – 55)	

Keterangan :
** Uji t tidak berpasangan*
*** Uji Mann Whitney*
**** Uji Kruskal-Wallis, Uji Post hoc Mann-Whitney : Jarang vs Sering p < 0,001, Jarang vs Selalu p < 0,001, Sering vs Selalu p < 0,001*

Diskusi

Derajat kesesuaian antar-penguji yang baik dalam menilai keterampilan klinik menandakan bahwa para penguji mempunyai objektivitas yang baik dalam menilai keterampilan klinik mahasiswa, sehingga dapat menghasilkan mahasiswa yang kompeten dalam melakukan tindakan keterampilan klinis. Selain itu derajat kesesuaian antar-penguji dapat mempengaruhi reliabilitas ujian OSCE.[13]

Derajat kesesuaian antar-penguji yang tidak baik dapat disebabkan oleh beberapa faktor, yaitu : objektifitas alat uji, tingkat kesulitan soal yang diujikan, penjelasan deskripsi penilaian pada lembar penilaian tidak jelas sehingga menyebabkan perbedaan persepsi penguji pada saat memberikan nilai, pelatihan penguji yang tidak memadai, penguji yang tidak homogen baik dari segi pengalaman menguji; pengalaman menjadi pengajar keterampilan klinik;

pengalaman profesi, ketidakmampuan penguji dalam menginternalisasikan deskripsi penilaian, fluktuasi emosional, perubahan motivasi pada saat menguji, kehilangan perhatian pada saat menguji, dan kondisi fisik penguji.

Derajat kesesuaian antar-penguji OSCE yang kurang baik pada Program Studi D3 Kebidanan FK Unpad dapat disebabkan oleh beberapa faktor, yaitu : 1) seluruh penguji OSCE di Program Studi D3 Kebidanan FK Unpad belum mendapatkan standardisasi atau pelatihan penguji OSCE yang memadai walaupun di Program Studi ini ujian OSCE sudah dilaksanakan selama 5 tahun, padahal menurut Wang pelatihan penguji OSCE sangat penting dan dapat membuat derajat kesuaian antar-penguji OSCE menjadi baik,[15] 2) penguji OSCE yang tidak homogen pada setiap pojok uji diantaranya adalah para penguji OSCE ini tidak seluruhnya mempunyai pengalaman menguji yang sama karena semakin sering menguji OSCE, penguji akan mempunyai pengalaman banyak dalam menguji sehingga dapat menghasilkan penilaian yang baik pula; belum seluruh penguji mengikuti pelatihan mengajarkan keterampilan seperti CTS dan TOT, pelatihan ini sangat penting dalam menunjang cara mengajar keterampilan klinik yang baik kepada mahasiswa dan meningkatkan pemahaman penguji dalam melakukan keterampilan klinik, 3) penguji dalam setiap pojok uji tidak seluruhnya menjadi pengajar keterampilan klinik yang diujikan, jika penguji memberikan penilaian pada keterampilan klinis yang diajarkannya maka penguji akan memberikan nilai lebih objektif karena penguji lebih paham akan materi yang diujikan, 4) penguji juga tidak seluruhnya mempunyai pengalaman klinik (tidak seluruh dosen di Program Studi D3 Kebidanan FK Unpad aktif berpraktik sebagai bidan diluar tugasnya sebagai dosen), jika penguji mengaplikasikan keterampilannya dalam praktik maka penguji lebih paham akan keterampilan klinis yang diujikan sehingga dapat memberikan penilaian yang lebih baik.

Selain faktor penguji, faktor objektivitas alat uji dapat mempengaruhi derajat kesesuaian penilaian antar-penguji karena lembar penilaian yang rinci dan yang mudah dipahami akan memudahkan penguji dalam menilai keterampilan. Pada Program D3 Kebidanan FK Unpad objektivitas alat uji belum dapat diketahui apakah menjadi salah satu penyebab kurang baiknya derajat kesesuaian antar-penguji OSCE dikarenakan alat uji tersebut belum pernah dianalisis reliabilitasnya.

Untuk dapat meningkatkan derajat kesesuaian antar-penguji dapat dilakukan berbagai cara yaitu : 1) melakukan pelatihan penguji OSCE dimana para penguji melakukan menilai keterampilan klinik mahasiswa, penilaian ini dapat menggunakan video atau simulasi ujian lalu para penguji mendiskusikan hasil penilaian tersebut terutama jika para penguji memberikan hasil penilaian yang berbeda dengan cara para penguji mendiskusikan rasionalisasi penilaiannya sehingga akhirnya didapatkan standar penilaian sehingga seluruh penguji mempunyai persepsi yang sama dalam menguji; 2) melakukan penyegaran pelatihan penguji OSCE dalam jangka waktu tertentu. Dengan berkembangnya berbagai macam penelitian tentang penanganan suatu penyakit, tentu akan ada perubahan dalam teknik keterampilan klinis kebidanan atau perubahan dalam sistem penilaian keterampilan klinis sehingga para penguji OSCE harus mengikuti penyegaran pelatihan penguji agar dapat memberikan penilaian ujian yang objektif ; 3) membuat lembar penilaian yang lebih rinci dan mudah dipahami oleh seluruh penguji. [13,15,16]

Simpulan

Agar derajat kesesuaian antar-penguji baik, maka syarat menjadi penguji OSCE harus terpenuhi.

Daftar Pustaka

1. Tambunan T, Soetjiningsih, Supriyatno B. Sistem Evaluasi pada Pendidikan Dokter Berbasis Kompetensi. Jakarta:

Kolegium Ilmu Kesehatan Anak Indonesia; 2011.
2. Amin Z, Eng KH. Basics in Medical Education. Singapore: World Scientific Publishing Co. Pte. Ltd.; 2003.
3. Epstein RM. Assessment in Medical Education. New England Journal of Medicine. 2007;356(4):387-96.
4. Turner JL, Dankoski ME. Objective structured clinical exams: a critical review. Family Medicine. 2008 Sep;40(8):574-8.
5. Miller JK. Competency-Based Training: Objective Structured Clinical Exercises (OSCE) in Marriage and Family Therapy. Journal of Marital and Family Therapy. 2010;36(3):320-32.
6. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. British Medical Journal. 1975 1975-02-22 1(5955):447-51.
7. Caballero C, Creed F, Gochmanski C, Lovegrove J. Nursing OSCEs A Complete Guide to Exam Success. New York: Oxford University Press; 2012.
8. Newble D. Techniques for measuring clinical competence: objective structured clinical examinations. Medical Education. 2004;38(2):199-203. Diunduh tanggal 14/05/2012.
9. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. Academic Medicine. 1998;73(9):993-7.
10. Project H. Pedoman Persiapan dan Penyelengaraan Objective Structured Clinical Examination (OSCE) Uji Kompetensi Dokter dan Dokter Gigi Indonesia. RI KPdK, editor. Jakarta2011.
11. Reznick R, Smee S, Rothman A, Chalmers A, Swanson D, Dufresne L, et al. An objective structured clinical examination for the licentiate: report of the pilot project of the Medical Council of Canada. Academic Medicine. 1992;67(8):487-94. Diunduh tanggal 23/05/2012.
12. Affandi Y. OSCE UKDI : Antara 'Outcome' Daya Ingat 90% dan Keraguan 52% Mahasiswa2012; RUBRIK OPINI – 3rd Spektrum BPN ISMKI: Available from: http://www.bpn.ismki.org/osce-ukdi-antara-outcome-daya-ingat-90-dan-keraguan-52-mahasiswa-2/.
13. Wilkinson TJ, Frampton CM, Thompson-Fawcett M, Egan T. Objectivity in Objective Structured Clinical Examinations: Checklists Are No Substitute for Examiner Commitment. Academic Medicine. 2003;78(2):219-23.
14. Hodges B, Regehr G, McNaughton N, Tiberius R, Hanson M. OSCE checklists do not capture increasing levels of expertise. Academic medicine : Journal of the Association of American Medical Colleges. 1999;74(10):1129-34.
15. Bresciani MJ, Oakleaf M, Kolkhorst F, Nebeker C, Barlow J, Duncan K, et al. Examining Design and Inter-Rater Reliability of a Rubric Measuring Research Quality across Multiple Disciplines. Practical Assessment, Research and Evaluation. 2009;14(12):1-7.
16. Wang P. The Inter-rater Reliability in Scoring Composition. CCSE English Language Teaching. 2009 September 2009;2(3):39-43.