

Interrater Reliability dari Checklist OSCE Keterampilan Mencuci Tangan dan Memakai Sarung Tangan di Program Studi Ilmu Keperawatan UGM

Interrater Reliability of Hand Washing and Gloving Skills of OSCE Checklist at UGM Nursing Science Study Program

Nur Fajriyah Rosyidah¹, Eri Yanuar Akhmad Budi Sunaryo², Totok Harjanto^{3*}

¹Program Studi Ilmu Keperawatan Fakultas Kedokteran, Kesehatan Masyarakat, dan Keperawatan Universitas Gadjah Mada

²Rumah Sakit Akademik Universitas Gadjah Mada

³Departemen Keperawatan Dasar dan Emergensi Fakultas Kedokteran Kesehatan Masyarakat dan Keperawatan Universitas Gadjah Mada

ABSTRACT

Background: Objective Structured Clinical Examination (OSCE) is a method to evaluate students' nursing skills. OSCE uses checklist as an instrument to test hand washing and gloving skills. While reliability values of both checklists are vital aspect for the instrument, they have not been measured in PSIK FKKMK UGM.

Objective: To identify the interrater reliability in the hand washing and gloving skills of OSCE checklists at PSIK FKKMK UGM.

Method: This research used non-experimental descriptive quantitative research type with cross-sectional design. The respondents were 92 first-year students at PSIK FKKMK UGM participating in the OSCE examination. Assessment on students performance were conducted using hand washing and gloving checklists by two raters. The scores were analyzed using Kappa and percent agreement (PA).

Result: Hand washing checklist had Kappa value of 0,146 (quite poor) and PA 80,40% (acceptable). Gloving checklist had Kappa value of 0,228 (poor) and PA 78,20% (acceptable). The results were divided into two categories, first Kappa not acceptable and PA acceptable or called Kappa paradox consisting of 14 checklist items. Furthermore, there were 11 checklist items in the second category namely Kappa not acceptable and PA not acceptable.

Conclusion: Hand washing and gloving skill checklists which are used by PSIK FKKMK UGM do not possess good interrater reliability in measuring the skills of nursing students. These checklists need revision and statistical test in order to improve education quality at School of Nursing at FKKMK UGM.

Keywords: checklist, gloving, hand washing, interrater reliability, OSCE

ABSTRAK

Latar belakang: Salah satu metode untuk mengevaluasi keterampilan keperawatan mahasiswa adalah menggunakan *Objective Structured Clinical Examination* (OSCE). Instrumen yang dapat digunakan dalam OSCE adalah *checklist*. Reliabilitas merupakan aspek penting dalam suatu instrumen. Pengujian reliabilitas dari *checklist* mencuci tangan dan memakai sarung tangan di PSIK FKKMK UGM belum pernah dilakukan.

Tujuan: Untuk mengetahui *interrater reliability* pada *checklist* OSCE keterampilan mencuci tangan dan memakai sarung tangan di PSIK FKKMK UGM.

Metode: Penelitian ini merupakan jenis penelitian deskriptif kuantitatif non-eksperimental dengan rancangan penelitian *cross-sectional*. Responden pada penelitian adalah 92 mahasiswa PSIK FKKMK UGM tahun pertama yang mengikuti ujian OSCE. Penilaian performa mahasiswa dilakukan menggunakan instrumen *checklist* mencuci tangan dan memakai sarung tangan oleh dua orang penguji. Hasil penelitian dianalisis menggunakan *Kappa* dan *percent agreement* (PA).

Hasil: *Checklist* mencuci tangan memiliki nilai *Kappa* 0,146 (cukup buruk) dan PA 80,40% (dapat diterima). *Checklist* memakai sarung tangan memiliki nilai *Kappa* 0,228 (buruk) dan PA 78,20% (dapat diterima). Terdapat 14-unit *checklist* yang masuk dalam kategori *Kappa* tidak dapat diterima, namun PA dapat diterima (paradoks *Kappa*). Terdapat 11-unit *checklist* yang masuk dalam kategori *Kappa* dan PA tidak dapat diterima.

Corresponding Author: Totok Harjanto

Gedung Ismangoen, Kompleks FKKMK UGM, Jl.Farmako Sekip Utara, Sendowo, Sinduadi, Mlati, Sleman, Yogyakarta 55281

Email: harjanto506@yahoo.com

Kesimpulan: *Checklist* keterampilan mencuci tangan dan memakai sarung tangan di PSIK FKKMK UGM belum memiliki *interrater reliability* yang baik dalam mengukur keterampilan mahasiswa. *Checklist* tersebut memerlukan perbaikan untuk menghasilkan reliabilitas baik demi kualitas pendidikan di Program Studi Ilmu Keperawatan FKKMK UGM yang lebih baik.

Kata kunci: *checklist*, cuci tangan, *interrater reliability*, OSCE, sarung tangan

PENDAHULUAN

Kunci pencegahan infeksi pada fasilitas pelayanan kesehatan adalah mengikuti prinsip pemeliharaan *hygiene* yang baik, disebut *standard precaution*. Dua contoh tindakan pencegahan infeksi yaitu dengan mencuci tangan dan menggunakan sarung tangan sebagai alat pelindung.¹ Sebelum melaksanakan praktik pelayanan keperawatan di RS pada tahap profesi, di tingkat akademik mahasiswa keperawatan mendapatkan pembelajaran mengenai keterampilan mencuci tangan dan memakai sarung tangan pada *skills laboratory* atau laboratorium keterampilan.² Salah satu institusi yang menggunakan ini adalah Program Studi Ilmu Keperawatan Fakultas Kedokteran Kesehatan Masyarakat dan Keperawatan UGM.

Untuk menjamin setiap perawat memiliki kompetensi yang dipersyaratkan, perlu dilakukan evaluasi sebelum melaksanakan praktik pelayanan keperawatan. Evaluasi untuk pembelajaran keterampilan yang diajarkan ketika di *skills laboratory* adalah dengan metode *Objective Structured Clinical Examination* (OSCE).³ Metode tersebut efektif untuk menilai praktik mahasiswa dalam hal keterampilan psikomotor dan pengetahuan teoretis terkait keterampilan yang diujikan.⁴ Instrumen yang digunakan dalam penilaian OSCE adalah *checklist*. Penggunaan instrumen *checklist* sebagai alat evaluasi OSCE ini perlu memperhatikan validitas dan reliabilitas.²

Mengukur reliabilitas dari suatu *checklist* dapat dilakukan dengan cara *interrater reliability*. *Interrater reliability* merupakan metode pengukuran yang dilakukan oleh dua pengukur yang berbeda.⁵ Pengukuran reliabilitas suatu instrumen penting untuk dilakukan karena jika alat ukur tidak reliabel, tidak dapat digunakan untuk mengetahui ciri atau keadaan sesungguhnya dari suatu objek ukur.⁶ Penelitian mengenai *interrater reliability* sudah pernah dilakukan. Namun, belum terdapat satu pun publikasi penelitian mengenai *interrater reliability* dari *checklist* OSCE mengenai keterampilan mencuci tangan dan memakai sarung tangan yang dimiliki oleh PSIK FKKMK UGM. Oleh karena itu, peneliti tertarik melakukan penelitian mengenai uji *interrater reliability* dari *checklist* keterampilan mencuci tangan dan memakai sarung tangan di PSIK FKKMK UGM sebagai upaya untuk mengetahui reliabilitas *checklist* tersebut.

Tujuan penelitian ini adalah untuk mengetahui *interrater reliability* pada OSCE keterampilan mencuci tangan dan memakai sarung tangan di PSIK FK-KMK UGM.

METODE PENELITIAN

Penelitian ini merupakan jenis penelitian deskriptif kuantitatif non-eksperimental dengan rancangan penelitian *cross-sectional*. Penelitian ini dilaksanakan di laboratorium keterampilan Gedung Ismangoen FKMK UGM. Responden penelitian ini adalah mahasiswa PSIK FKMK UGM tahun pertama yang berjumlah 92 mahasiswa. Pengambilan sampel menggunakan metode *total sampling* dengan kriteria yang mengikuti ujian OSCE keterampilan mencuci tangan dan memakai sarung tangan sesuai jadwal dan bersedia menjadi responden dalam penelitian.

Instrumen penelitian yang digunakan dalam penelitian ialah *checklist* keterampilan mencuci tangan dan memakai sarung tangan di PSIK FKMK UGM. *Checklist* yang digunakan sama antara proses *skills laboratory* dan saat OSCE. Terdapat 2 *checklist* dalam rangkaian keterampilan ini, pertama mengenai mencuci tangan dan kedua berupa *checklist* memakai sarung tangan. Di dalam *checklist* pertama terdapat 9 *item* yang terbagi atas dua tahap. Tahap pertama adalah tahap pra-interaksi yang berisi 1 *item*. Bobot maksimal yang diperoleh dari tahap pertama adalah 5. Selanjutnya pada tahap kedua yaitu tahap kerja yang berisi 7 *item* dengan bobot maksimal penilaian 95. Untuk *checklist* kedua ada 2 tahap, yaitu tahap pra-interaksi dan tahap kerja. Tahap kerja terbagi menjadi 2, pertama memakai sarung tangan dengan 6 *item* dan tahap kedua melepaskan sarung tangan dengan 4 *item* penilaian. Penilaian didasarkan pada 3 tingkatan yaitu, 0 jika *item* tidak dilakukan, 1 jika *item* dilakukan tetapi tidak sempurna, dan 2 jika *item* dilakukan dengan sempurna.

Pengumpulan data dilaksanakan secara langsung saat OSCE keterampilan mencuci tangan dan memakai sarung tangan yang berlangsung pada tanggal 28 - 29 Desember 2017 dan 2 Januari 2018. Setiap harinya terdapat 3 sesi dan satu sesinya diisi oleh 1 kelompok yang terdiri dari 10-11 mahasiswa. Kedua *rater* memberikan penilaian terhadap mahasiswa dengan format penilaian atau *checklist* yang telah disediakan.

Analisis data dilakukan pengelompokan hasil penilaian yang dilakukan oleh peneliti untuk mengetahui *Kappa* dan *percent agreement* keseluruhan *checklist* dan masing-masing *item*. Setelah itu dianalisis menggunakan uji *Kappa* dan *percent agreement* menggunakan perangkat lunak komputer. Nilai *Kappa* diinterpretasi berdasarkan Landis dan Koch.⁷ Untuk *interrater reliability* yang dapat diterima adalah apabila nilai *Kappa* $\geq 0,41$ dengan *PA* $>75\%$.^{8,9,10}

Surat persetujuan etik dikeluarkan oleh *Medical and Health Research Ethics Committee* (MHREC) Fakultas Kedokteran Kesehatan Masyarakat dan Keperawatan Universitas Gadjah Mada – Rumah Sakit Umum Pusat Dr. Sardjito dengan nomor referensi: KE/FK/1085/EC/2017. Setiap responden yang berpartisipasi dalam penelitian ini mendapatkan lembar penjelasan dan *informed consent* mengenai segala informasi yang berkaitan dengan penelitian. Data dan

informasi mengenai responden menjadi rahasia peneliti dan hanya dipakai untuk keperluan penelitian saja.

HASIL

Interrater reliability dari *checklist* keterampilan mencuci tangan dan memakai sarung tangan dilakukan dengan penghitungan *interrater reliability checklist* secara keseluruhan dengan mengelompokkan nilai mahasiswa menjadi dua kategori, yaitu kategori lulus dan tidak lulus. Kategori lulus dipenuhi apabila jumlah nilai dari *checklist* ≥ 75 . Sebaliknya, jika jumlah nilai yang didapat dari *checklist* < 75 , maka yang demikian termasuk dalam kategori tidak lulus. Kemudian dihitung jumlah kesepakatan antar-rater terhadap nilai lulus dan tidak lulus mahasiswa.

Hasil dari olah data untuk *interrater reliability* dari *checklist* keterampilan mencuci tangan, untuk penguji 1 terdapat 77 mahasiswa termasuk dalam kategori lulus dan 15 mahasiswa termasuk dalam kategori tidak lulus. Sementara untuk penguji 2, terdapat 83 mahasiswa termasuk dalam kategori lulus dan 9 mahasiswa termasuk dalam kategori tidak lulus. *Checklist* keterampilan mencuci tangan mempunyai nilai Kappa 0,146 dan nilai *percent agreement* (PA) 80,40%. Penilaian yang dilakukan oleh kedua penguji menghasilkan kesepakatan sejumlah 74 dan ketidaksepakatan sejumlah 18. Berdasarkan interpretasi Kappa menurut Landis dan Koch,⁷ nilai Kappa 0,146 termasuk dalam kategori cukup buruk yaitu berada pada rentang 0,00 – 0,20 sehingga *interrater reliability* dari *checklist* keterampilan mencuci tangan tidak dapat diterima. Interpretasi dari nilai PA juga dapat diterima (PA $< 85\%$).⁹ Hubungan antara hasil penilaian total *item checklist* keterampilan mencuci tangan dari kedua penguji disajikan pada Tabel 1.

Tabel 1. Nilai Kappa dan Percent Agreement (PA) Checklist Keterampilan Mencuci Tangan (n=92)

	Rater 2		Total	PA	Interpretasi PA	Kappa	Interpretasi (Landis dan Koch)
	Lulus	Tidak Lulus					
Rater 1	Lulus	71	6	77	80,4%	0,146	Cukup buruk
	Tidak Lulus	12	3	15			

Checklist kedua yaitu keterampilan memakai sarung tangan, pengelompokan mahasiswa lulus yaitu sebanyak 67 kali dan pengelompokan pada mahasiswa tidak lulus yaitu sebanyak 5 kali. Sementara, ketidaksepakatan ditemukan 4 kali dan 16 kali. Empat kali ketika *rater 1* mengelompokkan mahasiswa tidak lulus, namun *rater 2* mengelompokkan mahasiswa lulus dan 16 kali ketika *rater 1* mengelompokkan mahasiswa tidak lulus, namun *rater 2* mengelompokkan mahasiswa lulus. Nilai *Kappa* yang diperoleh dari hasil analisis data adalah 0,228 dan PA sebesar 78,20%. Berdasarkan kategori interpretasi nilai *Kappa* menurut Landis dan Koch⁷, interpretasi nilai *Kappa* dari *checklist* menggunakan sarung tangan tergolong

reliabilitas buruk yaitu pada rentang 0,21-0,40. Interpretasi PA masuk dalam kategori dapat diterima. Data selengkapnya dapat dilihat pada Tabel 2.

Tabel 2. Nilai Kappa dan Percent Agreement (PA) checklist keterampilan memakai sarung tangan (n=92)

	Rater 2		Total	PA	Interpretasi PA	Kappa	Interpretasi (Landis dan Koch)	
	Lulus	Tidak Lulus						
Rater 1	Lulus	67	16	83	78,2%	Dapat diterima	0,228	Buruk
	Tidak Lulus	4	5	9				

Tabel 3. Interpretasi interrater reliability item checklist mencuci tangan

Item	Nama Item dalam Checklist	Percent Agreement (%)	Nilai Kappa	Interpretasi Percent Agreement*	Kappa+
1	Menyiapkan alat.	89,1	-0,018	Dapat diterima	Tidak dapat diterima
2	Lepaskan semua perhiasan.	69,6	0,272	Tidak dapat diterima	Tidak dapat diterima
3	Nyalakan kran air.	80,4	0,065	Dapat diterima	Tidak dapat diterima
4	Basahi tangan dan pergelangan tangan dengan air. Pertahankan tangan lebih rendah dari siku untuk mengalirkan air melalui jari-jari tangan.	91,3	0,096	Dapat diterima	Tidak dapat diterima
5	Tuangkan sabun, ratakan ke semua bagian.	92,4	0,198	Dapat diterima	Tidak dapat diterima
6A	Telapak dengan telapak, gosoklah kedua telapak tangan secara bergantian sehingga kedua telapak tangan kena sabun.	83,7	0,048	Dapat diterima	Tidak dapat diterima
6B	Telapak tangan di atas punggung tangan kiri dan telapak kiri di atas punggung tangan kanan. Gosok kedua punggung tangan kanan secara bergantian dan gosok di antara jari jemari tangan secara bergantian sehingga kena sabun.	77,2	0,205	Dapat diterima	Tidak dapat diterima
6C	Telapak dengan telapak dan jari saling terkait. Gosok kedua telapak tangan di antara jari jemari secara bergantian sehingga kena sabun.	59,8	0,082	Tidak dapat diterima	Tidak dapat diterima
6D	Letakkan punggung jari telapak tangan satunya dengan jari saling mengunci. Gosok punggung jari yang saling mengunci pada telapak tangan satunya secara bergantian.	69,6	0,056	Tidak dapat diterima	Tidak dapat diterima
6E	Jempol kanan digosok memutar oleh telapak kiri dan sebaliknya. Gosok jempol dan jari-jari tangan lainnya secara memutar bergantian di kedua tangan.	76,1	0,222	Dapat diterima	Tidak dapat diterima
6F	Jari kiri menguncup dan gosok memutar ke kanan dan ke kiri pada telapak kanan dan sebaliknya. Gosok-gosoklah ujung kuku pada telapak tangan sehingga busa sabun masuk ke dalam sela-sela kuku secara bergantian di kedua tangan.	82,6	0,210	Dapat diterima	Tidak dapat diterima
7	Setelah selesai, siramlah kedua tangan dengan air yang mengalir dari kran.	81,5	0,070	Dapat diterima	Tidak dapat diterima
8	Keringkan kedua tangan dengan kain kering dan bersih.	68,5	0,075	Tidak dapat diterima	Tidak dapat diterima
9	Tutup kran dengan cara yang benar.	77,2	0,134	Dapat diterima	Tidak dapat diterima

Keterangan : *Kappa $\geq 0,41$ dapat diterima
+Percent Agreement $\geq 75\%$ dapat diterima

Interrater reliability dari setiap *item checklist* keterampilan mencuci tangan dapat dilihat pada Tabel 3. *Interrater reliability* dari *item checklist* keterampilan mencuci tangan menghasilkan rentang antara sangat buruk sampai buruk. Terdapat 1 *item* mempunyai interpretasi yang sangat buruk, 9 *item* dengan interpretasi cukup buruk dan 4 *item* mempunyai interpretasi buruk. Nilai *Kappa* tertinggi pada *checklist* yaitu 0,272 dan untuk nilai terendah yaitu -0,018. Sementara untuk interpretasi PA, dari 14 *item* terdapat 4 *item* mempunyai nilai PA yang tidak dapat diterima.

Interrater reliability pada setiap *item checklist* keterampilan memakai sarung tangan dapat dilihat di Tabel 4. *Interrater reliability* dari *item checklist* keterampilan memakai sarung tangan menghasilkan rentang antara sangat buruk sampai buruk. Terdapat 2 *item* mempunyai interpretasi yang sangat buruk, 7 *item* dengan interpretasi cukup buruk dan 2 *item* mempunyai interpretasi buruk. Nilai *Kappa* tertinggi pada *checklist* yaitu 0,292 dan untuk nilai terendah yaitu -0,062. Sementara untuk interpretasi PA terdapat terdapat 4 *item* PA yang dapat diterima dan 7 *item* mempunyai nilai PA yang tidak dapat diterima.

PEMBAHASAN

Instrumen keterampilan mencuci tangan dan memakai sarung tangan di PSIK FKMK UGM memiliki nilai *Kappa* dan *Percent Agreement* yang berbeda. Nilai *Kappa checklist* mencuci tangan yaitu 0,228 dan nilai *Kappa* untuk *checklist* memakai sarung tangan yaitu 0,146. *Checklist* memakai sarung tangan masuk dalam kategori buruk karena nilai *Kappa* berada antara 0,21 - 0,40. Untuk *checklist* mencuci tangan masuk dalam kategori cukup buruk karena nilai *Kappa* berada di antara 0,00 - 0,20.¹⁰

Nilai PA pada masing-masing *checklist* juga berbeda yaitu 80,40% dan 78,20% dengan interpretasi dapat diterima untuk *checklist* mencuci tangan dan *checklist* memakai sarung tangan. Nilai *Kappa* yang dapat diterima apabila mempunyai nilai $\geq 0,41$, sedangkan untuk PA yang diterima apabila persentasenya $\geq 75\%$. Jika dilihat dari parameter ahli tersebut, terdapat kontradiksi antara interpretasi nilai *Kappa* dan *Percent Agreement* pada *checklist* mencuci tangan dan memakai sarung tangan ini. Hasil penelitian menunjukkan bahwa di satu sisi *Kappa* bisa dikatakan buruk dan cukup buruk, sedangkan di sisi lain *Percent Agreement checklist* dapat diterima. Kategori dengan *Kappa* rendah, namun PA tinggi dapat disebabkan oleh sifat *Kappa* yaitu paradoks yang terjadi karena adanya pengaruh prevalensi dan efek bias.¹¹ Efek prevalensi pada paradoks *Kappa* berpengaruh pada ketidakseimbangan distribusi total marginal tabel 2x2 baik secara horizontal maupun vertikal yang akan memengaruhi nilai *Kappa* yang didapatkan.¹²

Jika dilihat keseluruhan penilaian yang dilakukan oleh kedua penguji, didapatkan hasil pengukuran *interrater reliability* memiliki reliabilitas yang rendah. Hal tersebut dapat terjadi dikarenakan beberapa faktor. Beberapa faktor tersebut adalah pelatihan *rater*, pengalaman

rater, jumlah kategori dalam instrumen, dan posisi rater terhadap posisi peserta dalam OSCE.¹³ Faktor lain yaitu adanya kelelahan dan faktor konsentrasi, baik dari rater maupun mahasiswa.¹⁴

Kelompok ini merupakan *item-item* yang memiliki nilai *Kappa* <0,41 dan *PA* \geq 75%. *Item-item* dari *checklist* mencuci tangan yang masuk dalam kelompok ini adalah *item* nomor 1, 3, 4, 5, 6A, 6B, 6E, 6F, 7 dan 9 (lihat tabel 3). Untuk *item-item* dari *checklist* memakai sarung tangan yang masuk dalam kelompok ini adalah *item* nomor 1, 2, 8 dan 9 (lihat tabel 4). Total seluruh *item* instrumen yang masuk dalam kelompok ini adalah 14 *item*. *Item-item* tersebut memiliki nilai *Kappa* yang rendah, namun kesepakatan antar penguji tinggi. *Kappa* tidak dapat diterima dan *PA* dapat diterima, terjadi ketika membandingkan koefisien secara komputasi dari kesepakatan yang telah diperoleh dari masing-masing penguji. Maka hasil yang ditunjukkan *PA* dapat diterima dengan perbandingan hasil yang signifikan dengan hasil *Kappa*.¹⁵ Kontradiksi hasil *Kappa* dan *PA* ini dikatakan sebagai suatu paradoks. Nilai *Kappa* yang rendah mungkin terjadi karena memang *Kappa* sangat bergantung pada nilai prevalensi, bias indeks, serta persebaran data yang tidak seimbang di tabel total marginal.¹⁶ Menggunakan nilai *Kappa* dan *PA* bersamaan akan mengurangi ambiguitas interpretasi *Kappa*.¹⁷ Instrumen *checklist* yang memiliki *item-item* dengan paradoks *Kappa* dapat dipengaruhi beberapa hal sehingga nilai *Kappa* rendah, namun nilai *Percent Agreement* tinggi dalam menentukan *interrater reliability* sehingga *item-item* tersebut dapat dikatakan reliabel.¹⁸

Item-item yang tidak dapat diterima baik *Kappa* maupun *PA*-nya adalah *item* nomor 2, 6C 6D dan *item* 4 untuk *checklist* mencuci tangan. Untuk *checklist* memakai sarung tangan *item-item* yang tidak dapat diterima baik *Kappa* maupun *PA* ialah *item* nomor 3, 4, 5, 6, 7, 10 dan *item* 11. Total *item* yang masuk kelompok *Kappa* tidak dapat diterima dan *PA* tidak dapat diterima adalah 11 *item*. *Item-item* ini dapat dikatakan bahwa bernilai *interrater reliability* yang buruk dari kedua segi analisis, sehingga perlu perhatian dan perbaikan *item*.

KESIMPULAN DAN SARAN

Hasil dan analisis data dari penelitian ini menunjukkan bahwa *interrater reliability checklist* keterampilan mencuci tangan dan memakai sarung tangan di PSIK FKKMK UGM belum memiliki reliabilitas yang baik dalam mengukur keterampilan mahasiswa.

Saran bagi program studi ilmu keperawatan, agar dalam melakukan evaluasi keterampilan mencuci tangan dan memakai sarung tangan pada mahasiswa menggunakan instrumen yang reliabel, penyamaan persepsi penguji dalam melakukan penilaian terhadap performa mahasiswa. Bagi peneliti selanjutnya, harapannya dapat melakukan penelitian terkait faktor-faktor lain yang bisa menyebabkan ketidaksesuaian antara *Kappa* dan *Percent Agreement* secara lebih mendalam, sehingga memengaruhi reliabilitas pada OSCE di Program Studi Ilmu Keperawatan FK-KMK UGM.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada semua partisipan yang sudah berpartisipasi dalam penelitian ini, khususnya kepada sponsor, responden, dan dosen penguji OSCE sehingga penelitian ini dapat berjalan dengan lancar.

DAFTAR PUSTAKA

1. Nurhesti PO, & Kustanti A. Surgical Asepsis (Teknik Tanpa Singgung, Scrubbing, Gowning, Glovng). In Haryani, K. Aulawi, H. S. Pangastuti, P. O. Nurhesti, Purwanta, & A. Kustanti, *Buku Skill Lab Sensasi dan Integritas Jaringan Blok 2.5* (p. 36). Yogyakarta: Fakultas Kedokteran UGM; 2017.
2. Peyre SE, Peyre CG, Hagen JA, & Sullivan ME. *Reliability of A Procedural Checklist as A High-Stakes Measurement of Advanced Technical Skill. The American Journal of Surgery.* 2010: 110-114.
3. Wimmers PF & Schauer GF. Validating OSCE Performance: The Impact of General Intelligence. *Health Professions Education*; 2017.
4. Rush S, Ooms A, Marks-Maran D, Firth T. Students Perceptions of Practice Assessment in The Skills Laboratory: An Evaluation Study of OSCAs with Immediate Feedback. *Nurse Education in Practice.* 2014;6 27-634.
5. Murti B. *Validitas dan Reliabilitas Pengukuran.* Matrikulasi Program Studi Doktorat Fakultas Kedokteran UNS. Surakarta; 2011.
6. Sugiyono. *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif dan R&D.* Bandung: Alfabeta; 2010.
7. Landis JR & Koch G. The Measurement of Observer Agreement for Categorical Data. *Biometrics.* 1977: 159-174.
8. Stemler SE, & Tsai J. *Best Practice in Interrater Reliability: Three Common Approaches.* Arizona: SAGE Publication; 2008.
9. Osborne J. *Best Practice in Quantitative Methods.* California: SAGE Publishing; 2008.
10. Graham M, Milanowski A, & Miller J. *Measuring and Promoting Inter-rater Agreement of Teacher and Principal Performance Ratings.* United State: Center for Educator Compensation Reform Westat; 2012.
11. Cargo M, Stankov I, Thomas J, Saini M, Rogers P, Mayo-Wilson E, & Hannes K. Development, Inter-Reter Reliability and Feasibility of A Checklist to Asses Implementation in Systematic Reviews: the Case of Provider-Based Prevention and Treatment Programs Targeting Children and Youth. *BMC Medical Research Methodology.* 2015:15-37.
12. Li, X. *Kappa- a critical review.* Uppsala: Department of Statistics Uppsala University. 2010.
13. Bao S, Howard N, Spielholz P, Silverstein B, & Polissar N. *Interrater Reliability of Posture Observation. Human Factors.* 2009: 292-309.
14. Clarke S, Rainey D, & Traynor M. Using The Objective Structured Clinical Examination (OSCE) to Assess Orthopaedic Clinical Skill for the Registered Nurse. *International Journal of Orthopaedic and Trauma Nursing.* 2011: 92-101.
15. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hroobartsson A, Streiner DL. Guidelines for Reporting Reliability and Agreement Studies (GRRAS). *International Journal of Nursing Studies.* 2011: 661-667.
16. Kuppens S, Holden G, Barker K, Rosenberg G. A Kappa-related Decision: K, Y, G, or AC1. *Social Work Research.* 2011;35(3):185-190.
17. Morris R, Macneela P, Scott A, Treacy P, Hyde A, O'Brien J, Drennan J. Ambiguites and Conflicting Result: The Limitation of The Kappa Statistics in Establishing The Interrater Reliability of The Irish Nursing Minimum Data Set for Mental Health: A Discussion Paper. 2008:645-647.
18. Bajpai R. Evaluation of Inter-Rater Agreement and Inter-Rater Reliability for Observational Data: An Overview of Concepts and Methods. *Journal of the Indian Academy of Applied Psychology.* 2015; 41(3): 20-27.