



L'évaluation clinique objective en stations multiples en nutrition : Analyse docimologique

The objective clinical evaluation in multiple stations in nutrition : Docrimological analysis

Faten Mahjoub¹, Ramla Mizouri¹, Nadia Ben Amor¹, Kamilia Ounaissa², Wafa Achour³, Henda Jamoussi¹

1- Institut National de Nutrition de Tunis, service A -Faculté de médecine de Tunis,

2- Institut National de Nutrition de Tunis, Hôpital du jour / Faculté de médecine de Tunis,

3- Centre national de greffe de moelle osseuse de Tunis / Faculté de médecine de Tunis,

RÉSUMÉ

Contexte : Dans le cadre de sa réforme pédagogique, la Faculté de Médecine de Tunis a intégré depuis 1998 l'évaluation clinique objective en stations multiples (ECOSM), comme évaluation sanctionnelle à la fin des stages. Cependant, cette méthode n'a jamais été évaluée en nutrition.

Objectif : Réaliser une analyse docimologique de l'ECOSM en nutrition destinée aux étudiants de la 2ème année du Deuxième Cycle des Etudes Médicales (DCEM 2).

Méthodes : Il s'agissait d'une étude prospective descriptive transversale portant sur l'ECOSM à but sanctionnel, des étudiants de la DCEM 2 ayant effectué un stage à l'Institut National de Nutrition de Tunis au cours de l'année universitaire 2017-2018.

Résultats : La moyenne générale de l'ECOSM pour la population étudiée était de $13,92 \pm 1,45$ sur 20. La moitié des stations était de difficulté acceptable. Plus que le tiers des questions (37%) avaient une discrimination bonne ou excellente. Pour toutes les stations confondues, le coefficient alpha de Cronbach moyen était de 0,62 témoignant d'une homogénéité interne non acceptable avec une grande hétérogénéité.

De même, les coefficients alpha de Cronbach calculés pour toutes les stations et toutes les questions étaient strictement inférieurs à 0,7. Le coefficient d'aplatissement était négatif, égal à -0,946 et la courbe de distribution était platykurtique.

Conclusions : A l'issue de ce travail, nous recommandons de sensibiliser particulièrement les étudiants et les enseignants à l'importance de l'évaluation des différents objectifs au cours du stage avant l'épreuve de l'ECOSM afin de combler les lacunes et de s'assurer de l'atteinte des objectifs du stage.

Mots-clés : Pédagogie – Évaluation – Docimologie.

SUMMARY

Context: As part of its pedagogical reform, the Faculty of Medicine of Tunis integrated since 1998 the objective clinical evaluation in multiple stations (ECOSM), as sanctioning evaluation at the end of the internships. However, this method has never been evaluated in nutrition.

Aim: Carry out a docimological analysis of the ECOSM in nutrition intended for students of the 2nd year of the Second Cycle of Medical Studies (DCEM 2).

Methods: This was a transversal descriptive prospective study on the ECOSM for sanctioning purposes, students of DCEM 2 having carried out an internship at the National Institute of Nutrition in Tunis during the academic year 2017- 2018.

Results: The general average of the ECOSM for the population studied was 13.92 ± 1.45 out of 20. Half of the stations were of acceptable difficulty. More than a third of the questions (37%) had good or excellent discrimination. For all the stations combined, the average Cronbach's alpha coefficient was 0.62, indicating an unacceptable internal homogeneity with great heterogeneity. Likewise, the Cronbach's alpha coefficients calculated for all stations and all questions were strictly less than 0.7. The flattening coefficient was negative, equal to -0.946, and the distribution curve was platykurtic.

Conclusions: At the end of this work, we recommend making students and teachers particularly aware of the importance of evaluating the various objectives during the internship before the ECOSM test in order to fill in the gaps and ensure that the internship objectives are met.

Keywords: Pedagogy - Evaluation – Docimology.

Correspondance

Ramla Mizouri

Institut National de Nutrition de Tunis, service A / Faculté de médecine de Tunis,

E-mail : mizouriramla@gmail.com

INTRODUCTION

Au cours de leurs cursus, les étudiants en médecine doivent acquérir outre les habiletés techniques, des habiletés de recueil de données, relationnelles et de résolution de problèmes aussi bien diagnostiques que thérapeutiques. L'ensemble constitue la compétence clinique recherchée (1). Les stages cliniques du deuxième cycle des études médicales représentent une étape fondamentale de l'apprentissage des compétences cliniques.

Il est bien connu que la motivation des étudiants pour apprendre dépend étroitement de la qualité de l'évaluation (2). L'évaluation de la performance et de la formation des étudiants se doit d'être de grande qualité afin d'assurer la crédibilité de la profession médicale.

L'examen clinique objectif structuré (ECOS), est considéré comme « le gold standard » pour évaluer les étudiants en fin de stages (3). En effet, c'est l'instrument qui se rapproche le plus de l'évaluation idéale de la compétence clinique puisqu'il tient compte de son caractère multidimensionnel. L'ECOS permet d'évaluer des compétences complexes, des savoir-faire et savoir-être, par des mises en situations cliniques.

Dans le cadre de sa réforme pédagogique, la Faculté de Médecine de Tunis (FMT) a intégré depuis 1998 l'ECOS, appelé l'évaluation clinique objective en stations multiples (ECOSM), comme évaluation sanctionnelle à la fin des stages. Cependant, cette méthode n'a jamais été évaluée en Nutrition.

L'objectif était de réaliser une analyse docimologique de l'ECOSM en Nutrition destinée aux étudiants de la 2^{ème} année du Deuxième Cycle des Études Médicales (DCEM 2).

MÉTHODES

Il s'agissait d'une étude prospective descriptive transversale portant sur l'ECOSM à but sanctionnel, des étudiants en DCEM 2 ayant effectué un stage de cinq semaines aux trois services de nutrition et maladies nutritionnelles à l'Institut National de Nutrition de Tunis (INNT) lors du deuxième semestre de l'année universitaire 2017 - 2018. Nous avons inclus 34 étudiants qui ont été affectés par la FMT pour passer leur ECOSM à l'INNT le 3 Mai 2018. Les étudiants étaient répartis comme suit : 11 externes au service « A », 12 externes au service « C » et 11 externes à l'hôpital de jour. La population analysée a porté sur 33 étudiants (22 filles et 11 garçons) puisqu'un étudiant absent le jour de l'ECOSM a été exclu de l'étude.

Choix des objectifs à évaluer :

Le choix des thèmes des stations d'ECOSM a été effectué en se référant aux objectifs du stage de nutrition existant dans le carnet des stages cliniques du DCEM 2 de la FMT. Ces objectifs sont établis, en tenant compte de leur pertinence, par la section d'endocrino-nutrition selon le profil du médecin généraliste Tunisien.

Au total, seize objectifs du carnet de stage ont pu être évalués. Les thèmes des stations ont porté sur les trois domaines des compétences le savoir (n=4), le savoir-faire (n=7) et le savoir être (n=1).

Rédaction des stations :

Les stations ont été rédigées, dix jours avant l'épreuve. Le nombre de stations était déterminé en tenant compte de l'espace disponible et du nombre d'étudiants par service. L'épreuve a comporté 12 stations de cinq minutes chacune : six stations sans patient et sans observateur, cinq stations avec patient et observateur et une station avec observateur sans patient.

Correction de l'ECOSM :

Chaque évaluateur avait la charge de corriger les feuilles de réponses des étudiants affectés dans son service. Ces notes étaient additionnées à celles des stations avec observateur. Le score total des différentes stations était de 105. La note de chaque étudiant a été convertie sur 20. L'étudiant réussissait l'épreuve s'il obtenait une note ≥ 10 . D'autre part, afin d'évaluer la fiabilité de l'épreuve, toutes les copies ont été corrigées par les trois évaluateurs à distance de l'épreuve et une étude de concordance a été réalisée.

Analyse docimologique :

Nous avons utilisé pour l'analyse docimologique de l'épreuve et des différentes questions le fichier « Anltem.xls ». Il s'agit d'un fichier Excel, proposé et utilisé par le docimologue « Serge Normand », au bureau d'évaluation de la Faculté de Médecine de l'Université de Montréal à partir de l'an 2000 (4,5).

Indice de difficulté :

C'est le pourcentage de réussite à la question. L'indice de difficulté peut prendre des valeurs allant de 0 à 1.

L'interprétation de l'indice de difficulté est comme suit (6) :

Facile	Acceptable	Optimale	Difficile
$\geq 0,7$	0,3 - 0,7	0,5 - 0,6	$< 0,3$

Coefficient de discrimination :

Le coefficient de discrimination est un coefficient de corrélation entre les résultats obtenus à la station et ceux obtenus à l'ECOSM. Il varie de -1 à 1. L'interprétation des valeurs de la discrimination est comme suit (7) :

Discrimination excellente	Bonne discrimination	Discrimination moyenne	Mauvaise discrimination
$\geq 0,40$	0,30 - 0,39	0,20 - 0,29	$\leq 0,19$

Corrélation difficulté discrimination :

Elle a été calculée en utilisant le coefficient de corrélation de Pearson applicable pour étudier la corrélation entre deux variables numérique à variation linéaire (8,9).

Coefficient alpha de Cronbach :

Il permet d'évaluer la consistance ou l'homogénéité interne d'un ensemble de questions.

L'interprétation des valeurs de l'alpha de Cronbach est comme suit (10,11) :

- $\geq 0,90$: excellente homogénéité interne (à condition de vérifier que les questions ne sont pas similaires et que leur nombre n'est pas trop grand)
- $[0,8, 0,9[$: bonne ou satisfaisante homogénéité interne
- $[0,7, 0,8[$: homogénéité interne acceptable
- $< 0,70$: homogénéité interne non acceptable (grande hétérogénéité nécessitant la suppression de quelques questions)

Courbe de distribution des scores :

Dans notre étude, le chiffrier « AnItem.xls » utilisé nous a permis le traçage de la courbe de distribution des scores. L'analyse de la courbe de distribution des scores se base sur deux paramètres d'une grande importance : les coefficients d'asymétrie et d'aplatissement (12).

Coefficient d'asymétrie :

Le coefficient d'asymétrie a été obtenu sur la feuille « Distribution » à partir de la fonction Excel « COEFFICIENT.

ASYMETRIE ». Il renseigne sur la difficulté globale de l'épreuve.

❖ Coefficient d'asymétrie nul : la distribution est symétrique. Elle est normale (mode = moyenne = médiane). Les valeurs sont les mêmes de part et d'autre du centre de la distribution.

❖ Coefficient d'asymétrie positif : la courbe est à asymétrie positive (mode < médiane < moyenne). Elle est étalée à droite. C'est la représentation graphique d'un effet plancher : le test est difficile.

❖ Coefficient d'asymétrie négatif : la courbe est à asymétrie négative (moyenne < médiane < mode). Elle est étalée à gauche. C'est la représentation graphique d'un effet plafond : le test est facile.

Coefficient d'aplatissement :

Le coefficient d'aplatissement a été obtenu sur la feuille « Distribution » à partir de la fonction Excel « KURTOSIS ». Il mesure le degré d'aplatissement de la courbe qui permet de déterminer le groupe d'étudiants sur lequel porte la discrimination au plus (queues de la courbe ou son centre).

❖ Coefficient d'aplatissement négatif : la distribution est platykurtique. Elle est basse et relativement aplatie. Son pic est plus arrondi autour de la moyenne avec des extrémités plus courtes et resserrées où il est difficile de différencier les individus.

❖ Coefficient d'aplatissement positif : la distribution est leptokurtique. Elle est haute et mince. Elle est relativement pointue au niveau de la moyenne avec des extrémités plus longues et étendues. Dans ce cas, il est difficile de différencier les individus autour de la moyenne.

❖ Coefficient d'asymétrie égal à 0 : la distribution est mésokurtique. C'est une situation intermédiaire entre les deux.

Analyse statistique :

Les données ont été saisies à l'aide du logiciel Excel 2007 et analysées en utilisant le logiciel SPSS version 17.0. Les variables quantitatives ont été exprimées en moyenne \pm écart type.

Le test t de student a été utilisé pour la comparaison de deux moyennes sur échantillons indépendants.

Les liaisons entre 2 variables quantitatives ont été étudiées par le coefficient de corrélation de Pearson, et en cas de

non-validité par le coefficient de corrélation des rangs de Spearman.

La concordance entre 2 variables quantitatives mesurées avec la même méthode par plusieurs observateurs différents été évaluée par le coefficient de corrélation intraclasse.

Ce coefficient s'interprète comme suit :

concordance nulle ou très mauvaise	concordance médiocre	concordance moyenne	bonne concordance	très bonne concordance.
< 0,20	0,21-0,41	0,41-0,60	0,61-0,80	0,81 et plus

Dans toutes les comparaisons, le seuil de signification a été fixé à 5%.

Considérations éthiques

Tous les étudiants ont donné leur consentement oral pour la participation à l'étude après avoir été informés du sujet et des objectifs du travail.

RÉSULTATS

Moyenne générale des notes des étudiants à l'épreuve:

La moyenne générale de l'ECOSM pour la population étudiée était de $13,92 \pm 1,45$ sur 20 avec des extrêmes allant de 11,43 à 16,67.

Les moyennes des notes des stations par écrit (stations de 1 à 6) et celles avec observateur (stations de 7 à 12) étaient respectivement de $5,97 \pm 0,98$ et de $7,67 \pm 0,73$ avec une différence statistiquement significative ($p < 0,01$).

Analyse docimologique :

La moitié des stations était de difficulté acceptable (Figure 1).

Aucune question à indice égal à 0 ou 1 n'a été retrouvée. L'épreuve ne comportait aucune question difficile. Plus que la moitié des questions (53%) avaient un indice de difficulté facile. Le quart des questions (26%) étaient acceptables. Les questions de difficulté acceptable étaient notées dans 47% des cas.

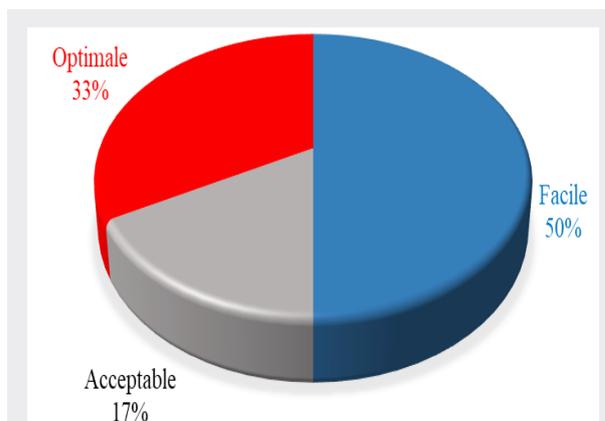


Figure 1. Répartition des stations en fonction de l'indice de difficulté.

Les stations à bonne ou excellente discrimination étaient notées dans 42% des cas (Figure 2).

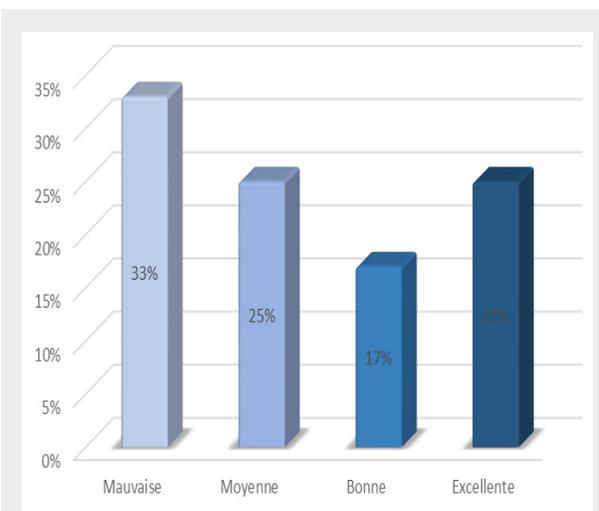


Figure 2. Répartition des stations en fonction de l'indice de discrimination.

Plus que le tiers des questions (37%) avaient une discrimination bonne ou excellente. Une discrimination moyenne était notée dans 21% des questions. Seules deux questions avaient un indice de discrimination négative.

Le coefficient de corrélation entre les indices moyens de difficulté et de discrimination de toutes les stations était de 0,131. Le coefficient de corrélation entre les indices moyens de difficulté et de discrimination de toutes les questions était de -0,245. La meilleure discrimination a été retrouvée pour les stations à difficulté variant de 0,4 à 0,6 (tableau 1).

Tableau 1. Coefficient de discrimination des différentes stations selon le niveau de difficulté.

Intervalles de difficulté	Indice de discrimination moyen
0,4-0,6	0,32
0,61-0,7	0,23
0,71-1	0,27

Une question idéale étant une question de difficulté acceptable et de bonne ou excellente discrimination, dans cette épreuve deux stations (1 et 5) et trois questions étaient idéales

(La 2^{ème} question de la première station et les 2 et 3^{èmes} questions de la station 6).

Pour toutes les stations confondues, le coefficient alpha de Cronbach moyen était de 0,62 témoignant d'une homogénéité interne non acceptable avec une grande hétérogénéité.

De même, les coefficients alpha de Cronbach calculés pour toutes les stations et toutes les questions étaient strictement inférieurs à 0,7.

Le tableau 2 résume les indices docimologiques des différentes stations et questions de l'ECOSM.

Le coefficient d'aplatissement était négatif, égal à -0,946. La courbe de distribution était platykurtique. Elle était basse et relativement aplatie. Il est difficile de différencier les individus. Le coefficient d'asymétrie était égal à 0,046, très proche de 0.

Une corrélation négative mais statistiquement non significative a été notée entre la note de l'ECOSM et le nombre d'absences au stage ($r = -0,33$; $P = 0,06$).

La concordance était très bonne entre les enseignants évaluateurs (tableau 3).

Tableau 2. Indices docimologiques des différentes stations et questions de l'épreuve.

		Indice de difficulté	Indice de discrimination	Alpha de Cronbach
Station 1		0,59	0,39	0,55
	Q1	0,81	-0,26	0,65
	Q2	0,50	0,58	0,53
Station 2		0,59	0,20	0,59
	Q1	0,88	0,09	0,62
	Q2	0,46	0,25	0,59
Station 3		0,42	0,16	0,60
	Q1	0,34	0,20	0,61
	Q2	0,47	0,16	0,61
Station 4		0,74	0,05	0,62
Station 5		0,60	0,68	0,54
Station 6		0,64	0,23	0,58
	Q1	0,82	0,23	0,65
	Q2	0,61	0,41	0,59
	Q3	0,54	0,36	0,59
Station 7		0,76	0,23	0,58
	Q1	0,65	0,18	0,60
	Q2	0,87	0,34	0,61
Station 8		0,80	0,38	0,57
	Q1	0,81	0,29	0,60
	Q2	0,77	0,22	0,61
Station 9		0,88	0,49	0,54
Station 10		0,83	0,02	0,62
Station 11		0,51	0,17	0,59
Station 12		0,83	0,44	0,55

* Q : question.

Tableau 3. Etude de la concordance des notes des stations écrites entre les évaluateurs.

Stations	Mesure de la concordance intraclasse
Station 1	0,9
Station 2	0,85
Station 3	0,93
Station 4	0,91
Station 5	0,93
Station 6	0,89

DISCUSSION

Dans la présente étude concernant 33 étudiants, la moyenne générale de l'ECOSM était de $13,92 \pm 1,45$ sur 20 avec des extrêmes allant de 11,43 à 16,67. La moitié des stations étaient de difficulté acceptable. Les stations avaient une discrimination bonne, moyenne ou excellente respectivement dans 17%, 25% et 25% des cas. Le coefficient alpha de Cronbach moyen était de 0,62 témoignant d'une grande hétérogénéité de l'épreuve. Le coefficient d'asymétrie était proche de zéro. L'épreuve n'était ni facile, ni difficile. Le coefficient d'aplatissement était négatif et la courbe de distribution était platykurtique. La discrimination était bonne entre les étudiants moyens mais faible entre les faibles et les forts. Les coefficients de concordance pour toutes les stations dépassaient 0,81 témoignant d'une très bonne concordance de notation entre les enseignants évaluateurs.

Ce Travail représente la première étude Tunisienne s'intéressant à l'ECOSM en Nutrition. Il a concerné tous les étudiants affectés aux trois services de l'INNT. Cette étude était basée sur l'analyse d'items (questions), technique la plus utilisée en docimologie. Ce qui a permis une comparaison avec les autres facultés de médecine. Néanmoins, l'effectif reste relativement faible ainsi que le nombre de stations.

Analyse des résultats de l'épreuve :

Tous les étudiants ont réussi à cette épreuve. Cependant, Le choix des stations était restreint aux objectifs enseignés aux trois services de l'INNT.

La moitié des stations était de difficulté acceptable dont 33% d'entre elles étaient optimales. Plus que la moitié des questions (53%) étaient faciles et 47% des questions étaient de difficulté acceptable. Des résultats similaires ont été notés dans une étude Tunisienne évaluant 32 étudiants par ECOSM (57% de questions faciles) (13).

La station 3 avait l'indice de difficulté le plus bas. Cette station est la seule comportant une notation négative. Aussi, elle a été introduite pour la première fois à l'ECOSM. La redondance d'une question dans différents examens la rend plus facile. L'indice de difficulté est influencé par plusieurs facteurs dont il faut tenir compte, notamment les caractéristiques du groupe d'étudiants ayant passé l'épreuve, leur niveau de compétence réelle, leur degré de préparation à l'évaluation ainsi que le facteur temps (14).

Le coefficient d'asymétrie était proche de zéro témoignant

d'un test ni facile ni difficile. Une épreuve de qualité en termes de difficulté, est une épreuve à difficulté acceptable. Ceci dit, l'épreuve doit comporter également des questions faciles permettant de distinguer les étudiants de faible niveau, ainsi que des questions difficiles permettant de différencier les étudiants de niveau de compétence plus élevé (9).

Le quart des stations avait une excellente discrimination alors que, 33% d'entre elles avaient une mauvaise discrimination. Dans les études de Baccouche (15) et Chahed (16) des stations à excellente discrimination ont été retrouvées, respectivement, dans 42% et 22% des cas. Un taux de 42% de questions de mauvaise discrimination a été noté. L'homogénéité du groupe d'étudiants, augmente le pouvoir discriminant des questions. Le coefficient d'aplatissement était négatif égal à -0,946. La courbe de distribution était platykurtique. La discrimination était bonne entre les étudiants moyens mais faible entre les faibles et les forts. Le but de l'ECOSM à la FMT n'est pas de discriminer entre les étudiants mais de valider le stage.

Le coefficient de corrélation entre les indices moyens de difficulté et de discrimination de toutes les stations était de 0,131. La meilleure discrimination a été retrouvée pour les stations à difficulté variant de 0,4 à 0,6. Bouzidi L et al. estiment que cet intervalle maximise la discrimination (8).

Dans cette épreuve, deux stations et trois questions étaient idéales. Dans une étude s'intéressant à l'ECOSM en Anatomopathologie, concernant 32 étudiants, une seule question idéale a été retrouvée (13).

Pour qu'un instrument d'évaluation soit efficace, il doit répondre à quatre critères ou qualités psychométriques à savoir la validité, l'objectivité, la faisabilité et la fiabilité (17-19).

Un instrument est fiable s'il mesure une variable donnée de façon constante et reproductible. Les principaux éléments de fiabilité sont la fiabilité inter-juges, qui mesure la concordance dans l'évaluation des performances des candidats par 2 juges différents ; et la fiabilité interne, méthode la plus employée, généralement mesurée par le coefficient α de Cronbach.

Celui-ci évalue jusqu'à quel point tous les items d'un test mesurent les mêmes caractéristiques ou les mêmes attributs (2, 17, 20).

Pour toutes les stations confondues, le coefficient

alpha de Cronbach moyen était de 0,62 témoignant d'une homogénéité interne non acceptable (grande hétérogénéité). De même, les coefficients alpha de Cronbach calculés pour toutes les stations et toutes les questions étaient $< 0,7$. Des résultats similaires ont été rapportés par Baccouche (15) avec un coefficient alpha de 0,61 de l'ECOSM en hématologie biologique. Une valeur nettement plus faible, de 0,29, a été rapportée en anatomopathologie (13). Muthusami, dans un travail analysant quatre études se rapportant à l'évaluation des résidents en chirurgie par ECOS, a rapporté que trois épreuves avaient une fiabilité élevée (coefficient $\alpha \geq 0,9$) alors que la 4^{ème} épreuve avait un coefficient α de 0,67 (21).

Dans la littérature, la valeur de cet indice oscille de 0,19 à 0,89. Brannick et al. en 2011, ont étudié 39 études contenant 188 valeurs du coefficient α (22). La valeur moyenne de cet indice était de 0,66 pour les stations et de 0,78 pour les questions. L'augmentation de la fiabilité était associée à une augmentation du nombre de stations et d'examineurs par station. Le nombre généralement admis est de l'ordre de 20 avec un minimum de 15 (2, 17, 19, 23).

Le coefficient alpha de Cronbach est l'outil statistique le plus puissant en matière d'homogénéité interne. Pour des examens type ECOS qui mesurent de multiples dimensions de la compétence clinique, des coefficients α de Cronbach supérieurs ou égaux à 0,6 sont considérés comme satisfaisants (17).

Pour estimer la fiabilité inter juges de l'ECOSM, on a eu recours à un test de concordance afin de comparer les notes administrées par les 3 évaluateurs. Les coefficients de concordance pour toutes les stations dépassaient 0,81 témoignant d'une très bonne concordance.

CONCLUSIONS

Les programmes d'enseignement médical visent à fournir des professionnels compétents, capables d'œuvrer dans le système de santé. Le choix de la méthode d'évaluation est crucial. En effet, les étudiants ajustent leur enseignement en fonction de leur évaluation.

A l'issue de ce travail, nous recommandons de sensibiliser particulièrement les étudiants et les enseignants à l'importance de l'évaluation des différents objectifs au cours du stage, d'augmenter le nombre de stations,

surtout celles évaluant le savoir-faire et le savoir être et d'élaborer des grilles de notation adaptées (grilles de correction et échelles d'évaluation) et validées et d'éviter la notation négative.

RÉFÉRENCES

1. Sibert L, Grand'Maison P, Charlin B, Grise P. Développement d'un Examen Clinique Objectif Structuré pour évaluer les compétences des internes en urologie. *Pédagogie Médicale* 2000;1(1):33-9.
2. Bernardin G, Mattéi M. Évaluation des stages hospitaliers par examen clinique objectif et structuré Bilan d'une expérience pédagogique au CHU de Nice. *Réan Urg* 1998;7(6):615-21.
3. Corsat M. Guide pour les examens des compétences cliniques (OSCE/ ECOS) A l'usage des enseignants évaluateurs [en ligne]. Disponible sur: <http://www.unifr.ch/clinical-skills/assets/files/Examens/guide%20eval.pdf>
4. Normand S. An Item, 2001 [en ligne]. Disponible sur : <http://www.medbev.umontreal.ca/docimo/cotonou.htm>.
5. Faculté de médecine Bureau d'évaluation Docimologie appliquée aux sciences de la santé, 2014 [en ligne]. Disponible sur : <https://www.slideserve.com/jonny/facult-de-m-decine-bureau-d-valuation-docimologie-appliqu-e-aux-sciences-de-la-sant>.
6. Hingorjo MR, Jaleel F. Analysis of One-Best MCQs: the Difficulty Index, Discrimination Index and Distractor Efficiency. *J Pak Med Assoc* 2012;62(2):6.
7. Donnon T, Lee M, Cairncross S. Using Item Analysis to Assess Objectively the Quality of the Calgary-Cambridge OSCE Checklist. *Can Med Educ J* 2011;2:16-22.
8. Bouzidi L, Jaillat A. Can Online Peer Assessment be Trusted? *J Educ Techno Soc* 2009;12:257-68.
9. Chartier P, Loarer E. L'analyse des items. Évaluer l'intelligence logique approche cognitive et dynamique. Paris: Dunod 2008:43-65.
10. Chiu J-M, Liu W-L. A Study of the Feasibility of Network Tutorial System in Taiwan. *J Educ Technol Soc* 2008;11(1):208-25.
11. Gliem JA, Gliem RR. Calculating, Interpreting and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales 2003, [en ligne]. Disponible sur: <https://scholarworks.iupui.edu/handle/1805/344>
12. Carricano M, Poujol F, Bertrandias L. Décrire les données. In: *Analyse de données avec SPSS*. Pearson. Paris; 2010. p. 33-8.
13. Limaiem F. Evaluation clinique objective en stations multiples en anatomie pathologique : analyse critique et étude de perception [Mémoire]. *Pédagogie Médicale*. Tunis; 2017.
14. Laveault D, Grégoire J. Introduction aux théories des tests en psychologie et en sciences de l'éducation, 2014 [en ligne]. Disponible sur: <https://eduq.info/xmlui/handle/11515/35674>

15. Baccouche H. Evaluation clinique objective en stations multiples en hématologie biologique : Analyse critique et docimologique [Mémoire]. Pédagogie Médicale. Tunis; 2017.
16. Chahed H. Evaluation clinique objective en stations multiples formative en Oto-Rhino-Laryngologie : Analyse critique [Mémoire]. Pédagogie Médicale. Tunis; 2017.
17. Sibert L, Grand'Maison P, Charlin B, Grise P. Evaluation de la compétence clinique en urologie : approche innovatrice basée sur l'observation de la performance. Progrès en urologie 1997;7:581-589.
18. Barman A. Critiques on the Objective Structured Clinical Examination. Ann Acad Med Singapore 2005;34:478-82.
19. Turner JL, Dankoski ME. Objective Structured Clinical Exams: A Critical Review. Fam Med 2008;40(8):574-8.
20. Wass V, Vander Vleuten C, Shatzer J, Jones R. Medical education quartet. Lancet 2001;357:945-49.
21. Muthusami A, Mohsina S, Sureshkumar S, Anandhi A, Elamurugan TP, Srinivasan K, et al. Efficacy and Feasibility of Objective Structured Clinical Examination in the Internal Assessment for Surgery Postgraduates. J Surg Educ 2017;74(3):398-405.
22. Brannick MT, Erol Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. Med Educ 2011;45(12):1181-9.
23. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011;2:53-5.