

La **C**ollaboration
nationale en matière
d'évaluation

Rapport
technique
annuel
2017



MEDICAL COUNCIL
OF CANADA

LE CONSEIL MÉDICAL
DU CANADA

TABLE DES MATIÈRES

APERÇU.....	4
1. ÉLABORATION DE L'EXAMEN	6
Plan directeur et spécifications des épreuves	6
Format de l'examen	8
Contenu de l'examen	8
Validité du contenu.....	9
Description des compétences et niveau de rendement attendu.....	11
2. DÉROULEMENT DE L'EXAMEN	13
Orientation des candidats	13
Centres d'examen et nombre de candidats	13
Personnel administratif de l'examen	14
Formation des patients standardisés (PS).....	14
Examineurs en chef (EC).....	15
Recrutement et formation des médecins examineurs (ME)	15
3. NOTATION DE L'EXAMEN	17
Assurance et contrôle de la qualité.....	17
Communication des résultats d'examen	18
Scores sur l'échelle des scores	18
Décision de réussite ou d'échec	23
Calcul des scores par compétence.....	23
4. RÉSULTATS DE L'ANALYSE PSYCHOMÉTRIQUE	24
Score sur l'échelle des scores	25
Statistiques pour les stations d'ECOS	28
Profil des résultats par compétence.....	31
Comparaisons historiques	35
5. RÉFÉRENCES BIBLIOGRAPHIQUES.....	36
ANNEXE A : RELEVÉ DE RÉSULTATS POUR L'EXAMEN DE LA CNE	37
ANNEXE B : RAPPORT D'INFORMATION SUPPLÉMENTAIRE POUR L'EXAMEN DE LA CNE	38

TABLEAUX ET FIGURES

Tableau 1 – Spécifications du contenu de l'examen de la CNE en 2017.....	7
Tableau 2 – Caractéristiques des stations d'ECOS de chacune des épreuves de l'examen de la CNE en 2017	10
Tableau 3 – Description des compétences de la CNE	12
Tableau 4 – Nombre de candidats à chaque épreuve de l'examen de la CNE en 2017.....	14
Figure 1 – Schéma d'ajustement statistique entre les versions d'épreuves de l'examen de la CNE en 2017	21
Tableau 5 – Statistiques globales des scores sur l'échelle de scores pour chacune des épreuves de 2017.....	25
Figure 2 – Distribution des notes par épreuve et pour toute l'année. Estimations de la fidélité des scores et décisions de classification.....	25
Tableau 6 – Estimation de fidélité, erreur type de mesure et indices de constance et d'exactitude des décisions pour chacune des épreuves de 2017	28
Tableau 7 – Statistiques globales sur les stations d'ECOS pour chacune des épreuves de 2017.....	30
Figure 3 – Profil des résultats par compétence à l'épreuve n° 1 de 2017 (tous les candidats).....	33
Figure 4 – Profil des résultats par compétence à l'épreuve n° 2 de 2017 (tous les candidats).....	33
Figure 5 – Profil des résultats par compétence à l'épreuve n° 3 de 2017 (tous les candidats).....	34
Figure 6 – Profil des résultats par compétence à l'épreuve n° 4 de 2017 (tous les candidats).....	34
Figure 7 – Profil des résultats par compétence à l'épreuve n° 5 de 2017 (tous les candidats).....	35
Tableau 8 – Taux de réussite des candidats de 2013 à 2017.....	35

APERÇU

Après la parution en 2004 du *Rapport du Groupe de travail sur le permis d'exercice des diplômés internationaux en médecine* (CCPSSRH, 2004), le Conseil médical du Canada (CMC) a entrepris une série d'initiatives connexes visant à appuyer l'évaluation et la formation des diplômés internationaux en médecine (DIM) au Canada. Un comité directeur a été mis sur pied et a travaillé de 2005 à 2009 à l'élaboration d'un cadre et d'une structure de gouvernance pour la Collaboration nationale en matière d'évaluation (CNE).

La CNE est le fruit d'un partenariat entre des organismes canadiens qui cherchent à simplifier le processus d'évaluation auquel doivent se soumettre les DIM souhaitant obtenir un permis d'exercice de la médecine au Canada. Un élément important du programme de la CNE est l'examen clinique objectif structuré (ECOS) pancanadien, appelé « examen de la CNE » (ou tout simplement « l'examen », dans le présent document). L'examen a pour but d'aider les programmes d'internat en milieu clinique des facultés de médecine canadiennes à sélectionner les DIM en vue de la première année de formation postdoctorale. Cet examen national vise à éviter le dédoublement des efforts par les programmes d'évaluation provinciaux des DIM. Les directeurs des programmes d'internat sont en mesure d'utiliser les résultats des candidats aux fins de sélection des DIM admis dans leurs programmes. En 2017, l'examen de la CNE a été offert en Alberta, en Colombie-Britannique, au Manitoba, en Nouvelle-Écosse, en Ontario et au Québec, et il était absolument obligatoire pour faire une demande d'internat au Service canadien de jumelage des résidents (CaRMS) dans toutes les provinces.

Le Comité central de coordination de la Collaboration nationale en matière d'évaluation (C³CNE) est chargé de superviser et de faire passer l'examen de la CNE. Il relève du Conseil de direction du CMC.

Travaillant sous l'autorité du C³CNE, le Comité d'examen de la CNE (CEC) supervise la création et la mise à jour du contenu de l'examen de la CNE. Le CEC veille à ce que tout ce contenu soit conforme au Plan directeur de l'examen de la CNE, et que les directives aux examinateurs en ce qui concerne la notation, ainsi que le contenu et le format de l'examen, respectent les lignes directrices de la CNE. De plus, le CEC approuve la communication des résultats, y compris les décisions sur les problèmes soulevés en matière de notation de l'examen et d'assurance de la qualité.

Des politiques et procédures ont été définies pour assurer la comparabilité des résultats d'une année à l'autre, une communication plus rapide des résultats, un contrôle uniforme de la qualité ainsi qu'une assurance de la qualité de toutes les épreuves, peu importe le lieu et la date où elles ont lieu. Dans cette optique, le CMC a mis sur pied et continue de mettre à jour un répertoire de procédures structurées. Celles-ci contribuent à la standardisation du déroulement de l'examen dans toutes les régions et tous les centres d'examen, et constituent le fondement des documents destinés aux patients standardisés (PS) et aux médecins examinateurs (ME). La standardisation des procédures est nécessaire pour étayer l'argument de validité selon lequel les écarts entre les notes obtenues par les candidats sont dus aux différences sur le plan des compétences, telles qu'elles sont évaluées par l'examen de la CNE, et non aux différences entre les centres. De plus, ces politiques et procédures sont nécessaires dans le cas de programmes d'évaluation de nombreux candidats, tels que celui de la CNE, où les sessions d'examen peuvent être géographiquement dispersées et dont les résultats doivent être comparables et d'une qualité uniforme.

Ce rapport résume les divers aspects du déroulement de l'examen ainsi que les principales propriétés psychométriques de toutes les épreuves de l'examen de la CNE qui ont été administrées en 2017.

1. ÉLABORATION DE L'EXAMEN

Ce chapitre porte sur la façon dont le contenu de l'examen est élaboré, le format de l'examen, le Plan directeur et les spécifications des épreuves, l'échelle de notation et les critères d'évaluation des compétences.

Plan directeur et spécifications des épreuves

Le Plan directeur et les spécifications des épreuves ont été élaborés et approuvés par le CEC en 2010. Les membres du CEC ont défini et précisé le contenu visé au cours d'une série de réunions tenues de 2009 à 2011.

Depuis 2011, le CEC a utilisé le plan directeur original pour toutes les épreuves à l'exception de l'épreuve des connaissances thérapeutiques; en 2013 la composante écrite prise en charge thérapeutique de la CNE a été supprimée du plan directeur et incorporée à l'ECOS pour créer une nouvelle section de compétences de prise en charge thérapeutique.

Des spécifications ont été élaborées pour l'examen de la CNE afin de satisfaire au Plan directeur et de faire en sorte qu'un contenu semblable soit évalué dans chaque épreuve. Le Plan directeur et les spécifications de l'examen ont été établis de telle sorte que les candidats soient évalués sur la base d'un contenu comparable d'une épreuve à l'autre. Chaque épreuve comporte un ensemble de cas d'ECOS répondant le mieux possible aux exigences des spécifications de l'examen.

Le **tableau 1** résume les spécifications de l'examen de la CNE et donne le sommaire du contenu requis et des compétences à évaluer au cours d'une épreuve. Ceci inclut les compétences cliniques, appareils et systèmes, disciplines et groupes d'âge des patients. Une nouvelle exigence portant sur le sexe des patients a été ajoutée, afin d'assurer une bonne répartition des patients selon le sexe dans l'ensemble des stations.

Tableau 1 – Spécifications du contenu de l'examen de la CNE en 2017

COMPÉTENCES CLINIQUES	Nombre recommandé	APPAREILS ET SYSTÈMES	Nombre recommandé
Prise d'anamnèse	7	Appareil respiratoire	Au moins 1
Examen physique	2-3	Appareil cardiovasculaire	Au moins 1
Anamnèse et examen physique combinés	Au plus 1	Appareil digestif	Au moins 1
Organisation	10	Système musculo-squelettique	2-3
Capacité de communiquer	10	Appareil génito-urinaire	
Maîtrise de la langue	10	Système endocrinien	
Diagnostic	Au moins 3	Système nerveux	2-3
Interprétation des données	Au moins 3	Santé mentale	
Analyses	Au moins 3	Appareil reproducteur	
Prise en charge thérapeutique	Jusqu'à 20 % des cas doivent se rapporter aux connaissances thérapeutiques	Plusieurs appareils et systèmes	
DISCIPLINES	Nombre recommandé	GROUPES D'ÂGE*	Nombre recommandé
Médecine	2-4	Nouveau-né	1-2
Chirurgie	2-4	Nourrisson (2-23 mois)	
Psychiatrie	1-2	Enfant d'âge préscolaire (2-5 ans)	
Obstétrique-gynécologie	1-2	Enfant d'âge scolaire (6-12 ans)	
Pédiatrie	1-2	Adolescent (13-18 ans)	1-2
Médecine gériatrique	1-2	Jeune adulte (19-44 ans)	4-5
		Adulte d'âge mûr (45-64 ans)	
		Aîné (65 ans et plus)	2-3
Soins d'urgence	1	* Âge du patient sur lequel porte le cas, pas nécessairement l'âge du patient standardisé	

SEXE*

La proportion de patients masculins/féminins ne doit pas être supérieure à 6:4.

* Le sexe du patient standardisé peut être différent de celui du patient réel.

Format de l'examen

En 2017, les épreuves d'examen de la CNE comptaient 10 stations d'ECOS opérationnelles d'une durée de 11 minutes. Chaque station a été conçue pour évaluer neuf compétences cliniques : analyses, capacité de communiquer, examen physique, diagnostic, interprétation des données, maîtrise de la langue, organisation, prise d'anamnèse, et prise en charge thérapeutique.

À chaque station d'ECOS, un PS simulait un cas clinique suivant un scénario précis, et le rendement de chaque candidat était évalué par un ME. Chaque station permettait de mesurer entre cinq et neuf compétences cliniques.

Des procédures standardisées, dont des protocoles de formation des ME et des PS ainsi que des procédures d'analyse des données, ont été suivies afin de s'assurer que les résultats de l'examen de la CNE soient comparables pour l'ensemble des épreuves et des candidats.

Contenu de l'examen

Le contenu de l'examen de la CNE est élaboré par un comité d'experts des disciplines cliniques, ainsi que d'experts en pédagogie médicale et en évaluation. En 2017, deux ateliers d'élaboration de contenu pour la CNE ont été tenus. Lors de ces ateliers, des cas ou des stations d'ECOS ont été rédigés, examinés par des pairs et approuvés pour leur mise à l'essai. Ces ateliers mettent l'accent sur le contenu requis pour répondre aux exigences du Plan directeur et des spécifications de l'examen de la CNE.

Une épreuve de l'examen de la CNE inclut des stations opérationnelles et des nouvelles stations mises à l'essai. Les notes attribuées aux candidats sont fondées uniquement sur les stations opérationnelles, chacune comportant plusieurs items. Les stations mises à l'essai ne sont pas identifiées comme telles, sont présentées dans des conditions standardisées, mais ne sont pas comptabilisées dans la note finale. Les données afférentes

à ces stations mises à l'essai sont examinées et analysées après l'examen. Les commentaires des ME, des PS et des formateurs de PS concernant ces nouvelles stations mises à l'essai sont également pris en considération, et pourraient faire au besoin l'objet d'ajustements. En cas d'ajustements importants, ces stations sont mises à l'essai une deuxième fois avant de devenir opérationnelles.

Pour faire en sorte que toutes les épreuves de la CNE soient comparables, chaque épreuve doit satisfaire à des critères précis (voir le **tableau 1 pour les spécifications de l'examen**). Les épreuves (pour chaque jour d'examen) sont proposées par l'agent d'élaboration d'examens, puis examinées, révisées et approuvées par le CEC.

Validité du contenu

Le degré de correspondance d'une épreuve par rapport aux spécifications du contenu de l'examen est une donnée probante qui appuie l'interprétation des scores à l'égard du but visé par l'examen (Kane, 2006; 2013). Cette section présente les spécifications des épreuves et mesure le degré de précision auquel chaque épreuve correspond ainsi que les compétences requises.

Le CEC collabore avec le personnel du CMC à la sélection et à l'approbation des cas d'ECOS pour une épreuve donnée. L'agent d'élaboration d'examens de la CNE prépare une ébauche des épreuves conformément aux spécifications du Plan directeur. Le CEC examine ensuite les épreuves, notamment chacune des stations d'ECOS, pour s'assurer qu'elles satisfont aux spécifications et que leur contenu se situe à un niveau de compétence approprié (celui d'un récent diplômé d'une faculté de médecine canadienne). Le CEC approuve la version finale du contenu de chaque épreuve. Pour des raisons de sécurité, plusieurs épreuves sont utilisées chaque année.

Le **tableau 2** montre l'échantillonnage des spécifications, des compétences cliniques et du nombre de stations pour chacune des 5 épreuves qui ont eu

lieu en 2017. La colonne *Nombre recommandé* indique le nombre souhaité de stations dans chaque épreuve pour chaque compétence clinique, discipline, sexe, appareil ou système, ainsi que pour chaque groupe d'âge. Une épreuve (n° 1) a été utilisée en mars, et quatre autres (n°s 2, 3, 4 et 5) en septembre.

Tableau 2 – Caractéristiques des stations d'ECOS de chacune des épreuves de l'examen de la CNE en 2017

		MARS	SEPTEMBRE			
COMPÉTENCES CLINIQUES	Nombre recommandé	Épr. n° 1	Épr. n° 2	Épr. n° 3	Épr. n° 4	Épr. n° 5
Prise d'anamnèse	7	7	7	7	7	7
Examen physique	2-3	2	2	2	2	2
Anamnèse et examen physique combinés	Au plus 1	1	1	1	1	1
Organisation	10	10	10	10	10	10
Capacité de communiquer	10	10	10	10	10	10
Maîtrise de la langue	10	10	10	10	10	10
Diagnostic	Au moins 3	8	8	9	8	9
Interprétation des données	Au moins 3	3	3	3	3	4
Analyses	Au moins 3	8	7	8	6	7
Thérapeutique et prise en charge	Au moins 3	10	9	8	9	8
DISCIPLINE*	Nombre recommandé	Épr. n° 1	Épr. n° 2	Épr. n° 3	Épr. n° 4	Épr. n° 5
Médecine	2-4	6	4	4	6	4
Chirurgie	2-4	3	2	4	3	2
Psychiatrie	1-2	1	2	1	1	1
Obstétrique-gynécologie	1-2	2	2	3	1	1
Pédiatrie	1-2	1	1	2	2	1
Médecine gériatrique	1-2	1	1	1	1	1
Soins d'urgence	1	3	3	1	1	3
SEXE**		Épr. n° 1	Épr. n° 2	Épr. n° 3	Épr. n° 4	Épr. n° 5
La proportion hommes/femmes ne devrait pas être supérieure à 60/40.		H = 5	H = 4	H = 6	H = 5	M = 4
		F = 5	F = 6	F = 4	F = 5	F = 6

* Le sexe du patient standardisé peut être différent de celui du patient réel.

Tableau 2 (suite) – Caractéristiques des ECOS de chacune des épreuves de l'examen de la CNE en 2017

APPAREILS et SYSTÈMES	Nombre recommandé	Épr. n° 1	Épr. n° 2	Épr. n° 3	Épr. n° 4	Épr. n° 5
Appareil respiratoire	Au moins 1	1	2	1	1	2
Appareil cardiovasculaire	Au moins 1	1	1	1	1	1
Appareil digestif	Au moins 1	1	1	1	1	3
Système musculo-squelettique	2-3	3	5	4	6	3
Appareil génito-urinaire						
Système endocrinien						
Système nerveux						
Santé mentale	2-3	5	3	5	4	3
Appareil reproducteur						
Plusieurs appareils et systèmes						
ÂGE*	Nombre recommandé	Épr. n° 1	Épr. n° 2	Épr. n° 3	Épr. n° 4	Épr. n° 5
Nouveau-né	1-2	1	1	2	1	1
Nourrisson (2-23 mois)						
Enfant d'âge préscolaire (2-5 ans)						
Enfant d'âge scolaire (6-12 ans)						
Adolescent (13-18 ans)	1-2	2	1	1	2	1
Jeune adulte (19-44 ans)	4-5	5	6	4	5	6
Adulte d'âge mûr (45-64 ans)						
Aîné (65 ans et plus)	2-3	2	2	2	2	2

* Âge du patient sur lequel porte le cas peut être différent de celui du patient standardisé.

Description des compétences et niveau de rendement attendu

En 2017, à l'aide d'échelles de notation (décrites ci-après), les ME ont noté le rendement des candidats pour chacune des compétences évaluées à chaque station d'ECOS. Un ME a été affecté à chaque station, si bien qu'à la

fin de l'ECOS, chaque candidat a été évalué par 12 ME, à raison d'un seul ME par station. Les notes attribuées par les ME ont servi à calculer les résultats selon la méthode décrite au chapitre de la *Notation de l'examen*.

En 2017, les ME ont noté le rendement des candidats pour chaque compétence en fonction du niveau de rendement attendu de la part d'un *récent diplômé d'une faculté de médecine canadienne*, à l'aide d'une échelle de type Likert en cinq points. Le **tableau 3** décrit ces cinq points ainsi que le rendement acceptable pour chaque compétence. Les ME ont reçu des documents d'orientation et de formation qui donnent plus de précisions à ce sujet dans le contexte de chaque station.

Tableau 3 – Description des compétences de la CNE

INACCEPTABLE comparativement à un récent diplômé d'une faculté de médecine canadienne	CAS LIMITE INACCEPTABLE comparativement à un récent diplômé d'une faculté de médecine canadienne	CAS LIMITE ACCEPTABLE comparativement à un récent diplômé d'une faculté de médecine canadienne	ACCEPTABLE comparativement à un récent diplômé d'une faculté de médecine canadienne	SUPÉRIEUR au niveau attendu d'un récent diplômé d'une faculté de médecine canadienne
○	○	○	●	○

Voici les désignations des niveaux de rendement ACCEPTABLE par compétence.

PRISE D'ANAMNÈSE

Attentes : Obtient auprès du patient, d'un membre de la famille ou d'une autre personne une description des événements pertinents qui est logique sur le plan médical et chronologique. Obtient de l'information suffisamment détaillée pour définir clairement le ou les problèmes du patient.

EXAMEN PHYSIQUE

Attentes : Procède à un examen physique dans une séquence logique efficace, ce qui permet de déterminer la présence ou l'absence d'anomalies et de corroborer la description du ou des problèmes du patient. Se préoccupe du confort et la pudeur du patient; explique ses gestes au patient.

ORGANISATION

Attentes : Approche l'interaction de façon cohérente, concise et méthodique. Gère son temps efficacement (y compris pendant les questions orales)

CAPACITÉ DE COMMUNIQUER

Attentes : Adopte une approche centrée sur le patient : inspire la confiance et le respect, et fait preuve de sensibilité. Fournit de l'information claire et s'assure de la compréhension du patient tout le long de la rencontre : encourage les questions et utilise la répétition et la récapitulation pour confirmer et/ou renforcer la compréhension. Respecte la confidentialité. Évite l'utilisation du jargon ou de termes trop familiers. Convient avec le patient, s'il y a lieu, d'un plan de prise en charge et de traitement. Tient un langage non verbal approprié (p. ex. contact visuel, gestuel, posture et pauses de silence).

MAÎTRISE DE LA LANGUE

Attentes : Parle clairement (volume et débit) avec une bonne prononciation; l'accent ne nuit pas à l'échange. Utilise une suite logique de mots et de phrases ainsi que les temps de verbe appropriés pour faire passer le message souhaité. Utilise le bon choix de mots et d'expressions dans le contexte donné (p. ex. vocabulaire).

DIAGNOSTIC

Attentes : Discerne l'information importante de l'information secondaire et parvient à un diagnostic différentiel et/ou un diagnostic raisonnable.

INTERPRÉTATION DES DONNÉES

Attentes : Interprète les données recueillies de manière appropriée dans le contexte donné.

ANALYSES

Attentes : Choisit les analyses de laboratoire et les examens diagnostiques appropriés pour poser ou confirmer le diagnostic; tient compte des risques et des bienfaits qui y sont associés.

THÉRAPEUTIQUE ET PRISE EN CHARGE

Attentes : Discute de la prise en charge thérapeutique, y compris, mais sans s'y limiter : pharmacothérapie*, effets indésirables et sécurité du patient, prévention et promotion d'un mode de vie sain, s'il y a lieu. Choisit les traitements appropriés (y compris surveillance, counseling, suivi); tient compte des risques et des bienfaits du traitement et en informe adéquatement le patient.

* Dans la plupart des cas, on s'attend à ce que les candidats désignent une classe de médicaments plutôt qu'un médicament spécifique (sauf, le cas échéant, dans les cas urgents), et l'on ne s'attend pas, généralement, à ce qu'ils indiquent les dosages exacts.

2. DÉROULEMENT DE L'EXAMEN

Ce chapitre décrit les moyens déployés pour assurer une standardisation adéquate du déroulement de l'examen, notamment par la tenue de séances d'orientation pour les candidats et de séances de formation pour les PS, ainsi que par l'établissement de critères de recrutement et la formation des ME.

Orientation des candidats

Le CMC fournit sur son site Web des renseignements détaillés à l'intention des candidats sur l'examen de la CNE : voir les sections « À quoi s'attendre le jour de l'examen », « Notation et résultats », ainsi que « Modalités d'inscription ».

Centres d'examen et nombre de candidats

Le **tableau 4** énumère les centres d'examen et le nombre de candidats pour chaque épreuve de la CNE en 2017.

Tableau 4 – Nombre de candidats à chaque épreuve de l'examen de la CNE en 2017

		Centres d'examen	Nbre total de candidats	Nbre de candidats pour la 1^{ère} fois	Nbre de candidats pour la 2^e fois ou +
<i>Mars</i>	Épr. n° 1	C.-B., Man., Ont.-Toronto, Qc	339	299	40
<i>Sept.</i>	Épr. n° 2	Alb.-Edmonton, N.-É., Ont.-Toronto	225	220	5
	Épr. n° 3	Alb.-Edmonton, C.-B., Ont.-Toronto	308	298	10
<i>Sept.</i>	Épr. n° 4	Alb.-Edmonton, Man., Ont.-Ottawa, Ont.-Toronto	382	375	7
	Épr. n° 5	Man., Ont.-Toronto, Qc	266	250	16
<i>Totaux</i>			1520	1442	78

Personnel administratif de l'examen

Chaque centre d'examen partenaire est responsable de l'embauche et de la supervision du personnel administratif. Les centres d'examen collaborent avec le CMC pour protéger les documents de l'examen et pour s'assurer que tous les intervenants (administrateurs des centres d'examen, formateurs des PS, PS, examinateurs en chef, ME, personnel présent lors des épreuves, traiteurs, etc.) collaborent congrûment. Les politiques et procédures de la CNE promulguées par le CMC assurent une standardisation adéquate du déroulement de l'examen. Les jours d'examen, le personnel du CMC supervise le personnel des centres d'examen partout au pays, que ce soit en personne, par téléphone ou par des moyens de communication électroniques, en plus d'offrir un service d'assistance téléphonique.

Formation des patients standardisés (PS)

Chaque centre d'examen est responsable de l'embauche et de la supervision des formateurs de PS, qui supervisent les PS et assurent la qualité et la

standardisation de leur performance lors des épreuves. Les PS sont formés dans chaque centre d'examen à l'aide des documents de formation de la CNE fournis par le CMC. Le personnel du CMC, notamment le coordonnateur de la formation de la CNE, fournit un soutien au chapitre de la formation des PS.

Examineurs en chef (EC)

Chaque centre d'examen de la CNE recrute un médecin comme EC. Son rôle dépend de la taille du centre d'examen et de la manière dont le gestionnaire du centre choisit de déléguer les tâches.

Chaque EC est chargé des tâches suivantes :

1. Aider, au besoin, à l'embauche et à la formation des ME;
2. Participer aux répétitions des PS avant les épreuves, et notamment faire une évaluation finale de la capacité des PS à agir le jour de l'examen de manière standardisée en suivant le scénario prescrit;
3. Superviser les examineurs et les candidats pendant l'examen;
4. Traiter, s'il y a lieu, les questions, les préoccupations et les plaintes des candidats le jour de l'examen;
5. Examiner et signer tous les rapports d'incident rédigés lors de l'examen.

Remarque : Un centre d'examen (celui de la Nouvelle-Écosse) recrute également un registraire adjoint, qui partage des responsabilités avec l'EC.

Recrutement et formation des médecins examineurs (ME)

Chaque ME de la CNE :

1. doit être licencié du Conseil médical du Canada (LCMC) et fournir son numéro de LCMC;

2. doit avoir une expérience récente de supervision d'externes ou de R-1, ou de l'expérience comme examinateur à ce niveau de formation;
 3. peut pratiquer en cabinet seulement (n'a pas besoin d'enseigner dans une université s'il satisfait à tous les autres critères);
 4. doit être de niveau R-4 ou plus OU détenir un certificat du Collège des médecins de famille du Canada (CMFC) s'il est médecin résident;
 5. doit avoir la capacité et l'endurance nécessaires pour accomplir la tâche; l'état de santé et l'âge pouvant être des facteurs critiques;
 6. peut ne pas être licencié du CMC s'il satisfait aux conditions suivantes :
 - être professeur à l'université;
- ET**
- être agréé par l'un des organismes ci-dessous et fournir son numéro d'agrément :
 - CRMCC
 - CMQ
 - CMFC

Avant chaque épreuve, tous les ME doivent participer à une séance d'orientation en ligne visant à mieux uniformiser la notation de l'examen. Cette séance en ligne comprend une séance de notation servant d'exercice et une discussion guidée. Le CMC offre aussi le jour de l'examen une séance d'orientation pour les ME. Cette séance est animée par l'EC et par le personnel des centres d'examen.

3. NOTATION DE L'EXAMEN

Ce chapitre décrit les procédures d'assurance et de contrôle de la qualité relatives à la notation de l'examen, ainsi que les procédures pour calculer les scores qui sont communiqués aux candidats et la manière dont ils sont calculés.

Assurance et contrôle de la qualité

Les opérations d'assurance de la qualité décrites ci-après sont effectuées pour garantir l'exactitude et l'intégrité des dossiers électroniques des candidats aux épreuves.

Les ME remplissent une feuille de notation pour chaque candidat qui se présente à leur station d'ECOS. Ces feuilles sont numérisées dans les centres d'examen et transmises de manière sécuritaire au CMC. Le personnel du CMC importe les données dans le logiciel Teleform^{MD} (un logiciel de Cardiff Software Inc) Les feuilles sont revues, les anomalies de numérisation identifiées (p. ex., un autocollant avec code à barres illisible, des annotations au crayon trop pâles d'un ME, etc.), et les corrections voulues sont effectuées dans les fichiers électroniques. Les données sont ensuite exportées électroniquement pour être traitées dans une application de notation, qui établit une liste de tous les candidats se situant à moins de 5 % au-dessus ou en dessous du seuil de réussite (approximativement une erreur type au dessus ou en dessous du seuil de réussite). Une fois les copies parvenues au CMC, les feuilles de tous les candidats qui figurent sur cette liste sont examinées par le personnel. Toutes les divergences entre les copies papier et les données électroniques sont corrigées dans les fichiers électroniques, afin que ceux-ci correspondent aux grilles de l'échelle de notation sur papier. Les fichiers électroniques mis à jour sont importés de nouveau dans l'application de notation, qui produit les relevés de résultats de tous les candidats.

Communication des résultats d'examen

Plusieurs éléments de l'examen de la CNE sont analysés et communiqués à chaque candidat : (1) score sur l'échelle des scores; (2) décision de réussite ou d'échec; (3) scores par compétence.

Les sections qui suivent résument les étapes de calcul des scores sur l'échelle des scores, de l'établissement du seuil de réussite, ainsi que l'estimation des scores par compétence à communiquer aux candidats, aux programmes destinés aux DIM et au Service canadien de jumelage des résidents (CaRMS).

Scores sur l'échelle des scores

De façon générale, le calcul des scores sur l'échelle des scores se fait en deux étapes. La première consiste à calculer le score total brut. La deuxième étape consiste à convertir les scores totaux bruts en scores sur l'échelle de scores finaux à communiquer aux candidats et aux programmes destinés aux DIM.

1ère étape : Calcul du score total

La 1ère opération de l'estimation du score total consiste à calculer la note pour chaque station d'ECOS, selon la formule suivante :

$$\text{Note pour la station} = \frac{[(\text{moyenne des échelles des notes obtenues sur les échelles de notation}) - 1]}{[(\text{note maximale des échelles de notation}) - 1]} * 100$$

où la moyenne des échelles de notation correspond à la moyenne des notes obtenues pour les compétences à une station donnée, et où la note maximale des échelles de notation (note maximale possible pour chaque échelle de notation) est de 5. Les soustractions indiquées dans le numérateur et le dénominateur visent à donner une fourchette de pourcentages correcte de 0 à 100. À titre d'exemple, dans le cas d'une

station portant sur 5 compétences pour lesquelles un candidat obtient respectivement des notes de 4, 4, 4, 5 et 5, la note totale calculée serait :

$$\begin{aligned} \text{note pour la station} &= \frac{((4 + 4 + 4 + 5 + 5)/5) - 1}{((5) - 1)} * 100 \\ &= \frac{((4,4) - 1)}{((5) - 1)} * 100 = 85,0 \end{aligned}$$

Les notes obtenues aux stations d'ECOS servent en premier lieu à calculer le score total de chaque candidat, qui correspond à la moyenne des notes obtenues aux dix stations d'ECOS. Ce score total est utilisé à la deuxième étape décrite plus loin.

Avant de décrire cette prochaine étape, il est important de mentionner que les calculs tiennent compte des données manquantes. Comme la note pour une station repose sur la moyenne des compétences évaluées pour cette station, il faut tenir compte des données manquantes de manière à ce qu'elles n'aient pas d'incidence négative sur la note d'un candidat. Une donnée est manquante lorsqu'un ME omet d'indiquer une note pour une compétence donnée sur la grille de l'échelle de notation numérisable. Lorsqu'un ME omet de fournir une note pour une échelle de notation, la note de la station est calculée selon la moyenne des notes fournies par le ME. Par exemple, dans le cas d'une station portant sur cinq compétences, mais où la note est manquante pour l'une d'entre elles, des notes moyennes de 4, 4, 4 et 5 pour les autres compétences donneraient la note ci-dessous pour la station :

$$\begin{aligned} \text{note pour la station} &= \frac{((4 + 4 + 4 + 5)/4) - 1}{((5) - 1)} * 100 \\ &= \frac{((4,25) - 1)}{((5) - 1)} * 100 = 81,25 \end{aligned}$$

Si la note manquante était traitée comme un zéro sans autre ajustement, la note totale pour cette station serait de 60 %.

2e étape : Ajustement statistique et scores sur l'échelle des scores

La 2^e étape, qui vise à établir le score sur l'échelle des scores, peut être subdivisée en deux volets. Dans un premier temps (2a), grâce aux stations communes, les notes obtenues aux épreuves de 2017 font l'objet d'un ajustement statistique pour être liées aux notes obtenues lors d'épreuves précédentes. Ce lien statistique remonte jusqu'à la 1^{ère} épreuve de 2013, qui a servi à déterminer le seuil de réussite et à définir l'échelle des scores. Ensuite (2b), les scores totaux ainsi ajustés sont convertis en scores sur l'échelle de scores à communiquer aux candidats et aux programmes destinés aux DIM.

2a : Ajustement statistique de scores

Chaque épreuve comporte un échantillon différent de stations, il se peut qu'une épreuve soit plus difficile qu'une autre.

Le processus d'ajustement statistique du score final tient compte du niveau légèrement variable de difficulté entre les épreuves et permet un ajustement des scores totaux bruts, afin que tous les scores finaux soient comparables sur l'échelle de scores. Cet ajustement statistique de scores fournit en outre un moyen d'appliquer le même seuil de rendement à des candidats qui passent des épreuves différentes. L'ouvrage de Kolen et Brennan intitulé *Test equating, scaling, and linking: Methods and Practice* (2^e édition, 2004) aborde en détail les schémas, les problèmes et les méthodes d'établissement d'un lien statistique. Ce processus permet de comparer au fil des ans les notes sur l'échelle des scores communiquées aux candidats et aux programmes destinés aux DIM.

Une méthode d'établissement d'un lien statistique entre des versions d'une épreuve, consiste à faire en sorte qu'une partie du contenu soit identique

d'une version d'une épreuve à une autre. Cette méthode est dite à items communs et à groupes non équivalents. On appelle ensemble de stations d'ancrage le sous-ensemble de stations qui est identique entre deux épreuves. Ces stations permettent d'obtenir une estimation statistique des compétences globales des candidats à chaque version de l'épreuve ainsi que du niveau de difficulté des versions de l'épreuve. Dans le cas des épreuves de la CNE en 2017, les cinq épreuves comportaient des ensembles de stations d'ancrage servant à faire un ajustement statistique du score final des candidats. En 2017, l'épreuve n° 1 a été choisie comme épreuve de base servant à faire l'ajustement statistique des scores finaux des candidats aux autres épreuves de 2017. La figure 1 illustre sous forme graphique le schéma d'ajustement statistique adopté. Elle montre que l'épreuve no 1 a servi d'épreuve de base pour effectuer l'ajustement statistique entre les stations d'ancrage. Chacune des autres épreuves reprenait environ 30 pour cent du contenu de l'épreuve n° 1. Cette partie du contenu a servi à établir un lien statistique entre les épreuves. La règle empirique veut qu'au moins 20 stations ou 20 pour cent de l'épreuve servent d'ancrage.

2017

MARS	SEPTEMBRE			
Épreuve n° 1	Épr. n° 2	Épr. n° 3	Épr. n° 4	Épr. n° 5
CONTENU D'ANCRAGE : 30 %	CONTENU D'ANCRAGE : 30 %			
CONTENU D'ANCRAGE : 30 %		CONTENU D'ANCRAGE : 30 %		
CONTENU D'ANCRAGE : 30 %			CONTENU D'ANCRAGE : 30 %	
CONTENU D'ANCRAGE : 30 %				CONTENU D'ANCRAGE : 30 %

Figure 1 – Schéma d'ajustement statistique entre les versions d'épreuves de l'examen de la CNE en 2017

La méthode de Levine pour les scores observés est utilisée pour ajuster les scores entre les diverses épreuves de l'examen de la CNE (Kolen et Brennan, 2004).

2b : Conversion en scores sur l'échelle des scores

Une fois les scores totaux bruts établis et ajustés à l'épreuve de base, ils sont convertis en scores sur l'échelle de scores destinées à être communiquées aux candidats. Une fonction linéaire, définie par une pente et une ordonnée à l'origine, est utilisée pour convertir tous les scores ajustés. Cette conversion finale permet de comparer directement par rapport à l'épreuve n° 1 de 2013 toute divergence observée dans la moyenne et l'écart-type des scores sur l'échelle de scores des épreuves de l'année en cours. Par exemple, une hausse moyenne de 70,0 à 75,0 ou une baisse de 70,0 à 65,0 indique que le niveau général des candidats ayant passé une épreuve subséquente est respectivement plus élevé ou plus faible. Dans le cas des épreuves nos 1 à 5, la formule de la conversion finale est la suivante :

$$\text{score sur l'échelle des scores}_x = (\text{pente})(\text{note ajustée}_x) + (\text{ordonnée à l'origine})$$

où

score sur l'échelle des scores_x est la fonction linéaire servant à calculer les notes sur l'échelle des scores pour le candidat X,
pente est égale à 1,06 et s'applique à toutes les versions d'épreuves (en fonction de la conversion initiale de l'épreuve n° 1 de 2013),
ordonnée à l'origine est égale à 3,34 et s'applique à toutes les versions d'épreuves (en fonction de la conversion initiale de l'épreuve n° 1 de 2013), et
note ajustée_x est le score lié à l'épreuve de base obtenue par le candidat X.

Après ces conversions pour les épreuves nos 1 à 5, tous les scores sur l'échelle des scores sont arrondis à un nombre entier entre 0 et 100 (à noter qu'il ne s'agit pas ici d'une échelle de pourcentages). Ce sont ces valeurs arrondies qui sont communiquées aux candidats et aux programmes destinés aux DIM. Par exemple, dans le cas d'une note ajustée de 81,25, la note sur l'échelle des scores serait :

$$\begin{aligned} & \text{note sur l'échelle des scores}_x \\ & = (1,06) * (81,25) + (3,34) = 89,47 \text{ arrondi à } 89 \end{aligned}$$

Décision de réussite ou d'échec

Le seuil de réussite pour cet examen a été déterminé par un comité de 18 médecins de partout au pays, représentant des facultés de médecine et différentes spécialités, et comptant plusieurs années d'expérience dans la supervision d'étudiants et de résidents. Les membres du comité ont revu le contenu des stations et des épreuves ainsi que les données sur les scores avant d'établir le seuil de réussite recommandé. L'épreuve n° 1 de mars 2013 a servi à établir le seuil de réussite à l'aide d'une méthode du groupe limite. Le document intitulé *Recommending a Passing Score on the National Assessment Collaboration (NAC) Examination: Procedures and results* (Hambleton et De Champlain, 2013) donne tous les détails au sujet du processus d'établissement du seuil de réussite. Les notes sur l'échelle des scores communiquées et le seuil de réussite de 65, établi et recommandé par le comité de médecins en 2013 et approuvé par le CEC en avril 2013, ont servi à attribuer à chaque candidat à l'examen en 2017 un résultat de réussite ou d'échec. Un score de 65 ou plus est considéré comme une réussite à l'examen.

Calcul des scores par compétence

Le calcul des scores par compétence est semblable à celui de la note à

attribuer pour une station donnée. On considère toutes les échelles de notation associées à une compétence donnée, puis on calcule la moyenne des échelles de notation pour cette compétence, que l'on convertit ensuite en une note de 0 à 100. À titre d'exemple, voici la formule générale employée dans le cas de la compétence « organisation » :

$$\text{score pour la compétence organisation} = \frac{[(\text{moyenne des notes obtenues sur les échelles de notation pour la compétence organisation}) - 1]}{[(\text{score maximal des échelles de notation}) - 1]} * 100$$

Les scores par compétence servent à établir les profils des résultats par compétence présentés sous forme graphique dans le Rapport d'information supplémentaire. Les scores par compétence ne servent pas à calculer le score total ou les scores sur l'échelle des scores (selon les méthodes décrites plus haut); par conséquent, on ne peut pas comparer directement les scores par compétence avec les scores finaux sur l'échelle des scores des candidats. Les scores par compétence fournissent aux candidats une indication générale de leur rendement à l'examen de la CNE.

4. RÉSULTATS DE L'ANALYSE PSYCHOMÉTRIQUE

Les sections qui suivent donnent des statistiques sommaires concernant les scores sur l'échelle de scores et les taux de réussite. Elles donnent aussi des estimations de la fidélité des scores et des décisions de classification, un résumé de la qualité des items, ainsi que le profil des résultats par compétence. Des résultats vérifiés et approuvés par le CEC pour les épreuves de mars et septembre sont utilisés dans cette section. Les candidats ayant reçu une mention « pas de résultat », « résultat refusé », ou qui ont manqué plus d'une station ne sont pas pris en compte.

Score sur l'échelle des scores

Le **tableau 5** présente des statistiques sommaires à propos des scores sur l'échelle des scores et des taux de réussite pour chacune des versions d'épreuves. Ces statistiques sont établies à partir des scores sur l'échelle des scores communiquées aux candidats. Le minimum, le maximum et l'écart type sont des indicateurs du degré de dispersion des scores sur l'échelle des scores.

Tableau 5 – Statistiques globales des scores sur l'échelle de scores pour chacune des épreuves de 2017

Session		N	Min.	Max.	Moyenne	Médiane	Écart-type	Taux de réussite
Mars	Épr. n° 1	330	46,0	91,0	73,7	74,5	7,5	87,9%
	Épr. n° 2	224	58,0	83,0	72,8	74,0	4,9	93,8 %
Sept.	Épr. n° 3	308	43,0	86,0	68,6	70,0	7,9	73,4 %
	Épr. n° 4	382	46,0	91,0	75,2	75,0	6,5	94,5 %
	Épr. n° 5	262	50,0	92,0	73,0	73,5	7,8	84,0 %
Totaux		1506	43,0	92,0	72,8	73,0	7,4	86,8 %

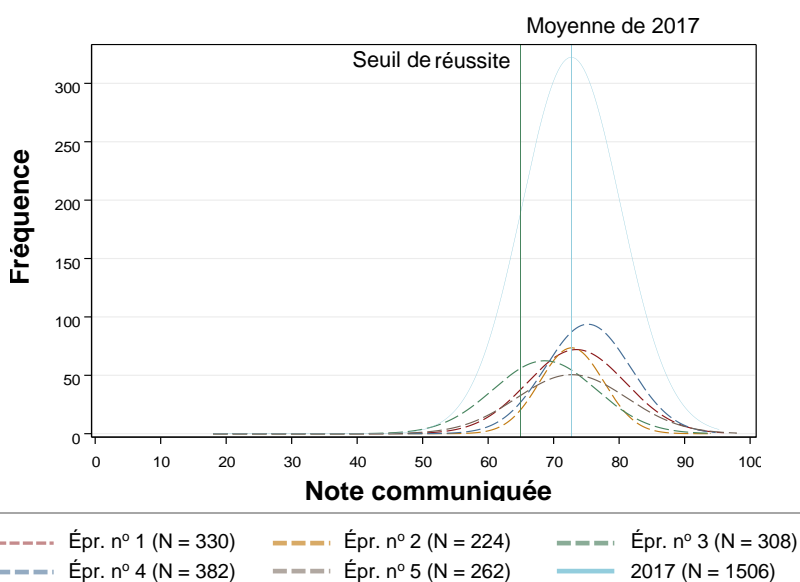


Figure 2 – Distribution des notes par épreuve et pour toute l'année. Estimations de la fidélité des scores et décisions de classification

Le tableau 6 montre les estimations de la fidélité des scores, l'erreur type de mesure (ETM), les indices de constance et d'exactitude des décisions, ainsi que les estimations correspondantes de faux positifs et de faux négatifs, pour chacune des versions d'épreuves. Les estimations de faux positifs indiquent la proportion de candidats ayant réussi selon leurs scores observés, mais qui auraient dû échouer selon leurs compétences réelles. Les estimations de faux négatifs indiquent la proportion de candidats qui ont échoué selon leurs scores observés, mais qui auraient dû réussir selon leurs compétences réelles.

Coefficient alpha de Cronbach

Le coefficient alpha de Cronbach a été utilisé pour estimer la fidélité des scores à l'examen de la CNE, c'est-à-dire la capacité de l'examen à évaluer correctement les effets des stations et le rendement des candidats. Cette estimation de la fidélité des scores est décrite plus en détail à la sous-section 2.4.4 du chapitre de Haertel sur la fidélité dans un ouvrage de Brennan intitulé Educational Measurement (Haertel, 2006). La formule du coefficient alpha de Cronbach est la suivante :

$$\alpha\rho_{XX'} = \frac{n}{n-1} \left(1 - \frac{\sum \sigma_{X_i}^2}{\sigma_X^2} \right)$$

où n représente le nombre de stations, $\sigma_{X_i}^2$ la variance pour chaque station, et σ_X^2 la variance du score total (Haertel, 2006, p. 74). L'estimation de fidélité des scores est un indicateur de la constance (ou reproductibilité) des résultats d'un examen sur plusieurs répétitions de mesure (Crocker et Algina, 1986; Haertel, 2006). Le **tableau 6** donne le coefficient alpha de Cronbach pour chaque épreuve. La combinaison de l'estimation de fidélité et de l'ETM totale de l'examen peut fournir une indication supplémentaire de la fidélité du score sur l'échelle des scores d'un candidat.

Erreur type de mesure

L'ETM indique jusqu'à quel point, dans les limites d'un certain intervalle de confiance (p. ex. 68 pour cent ou 95 pour cent), le score observé d'un candidat reflète le score vrai d'un candidat. L'ETM devrait avoir la valeur la plus petite possible, de sorte que l'évaluation des compétences du candidat soit aussi précise que possible. L'ETM se calcule comme suit :

$$ETM = \sigma_x \sqrt{1 - \alpha\rho_{XX'}}$$

où σ_x est l'écart-type du score total observé (racine carrée de la variance du score total), et $\alpha\rho_{XX'}$ l'estimation de la fidélité telle que décrite ci-dessus. Le tableau 6 donne l'ETM des notes sur l'échelle des scores communiquées pour chaque version d'une épreuve.

Consistance et exactitude des décisions

Les décisions de réussite ou d'échec sont cruciales dans un examen dont l'enjeu est aussi important que celui de la CNE. On peut aussi évaluer la fidélité de l'examen de la CNE en examinant la constance et l'exactitude des décisions de réussite ou d'échec à partir des scores de l'examen. On peut faire une estimation des indices de constance et d'exactitude des décisions à l'aide de la méthode de Livingston et Lewis (1995). L'indice de constance des décisions est une estimation de l'accord entre les décisions de réussite ou d'échec résultant d'épreuves ou de versions d'épreuves parallèles potentielles de l'examen. L'indice d'exactitude des décisions est une estimation de l'accord entre les décisions de réussite ou d'échec observées pour les candidats et celles qui seraient fondées sur leur score vrai (c'est-à-dire la note moyenne à laquelle on pourrait s'attendre si les candidats pouvaient être testés un nombre infini de fois).

Le **tableau 6** donne les indices de constance et d'exactitude des décisions pour chacune des épreuves. Idéalement, ces deux indices devraient être élevés, par exemple 0,80 ou plus, témoignant de décisions de réussite ou d'échec exactes et consistantes. Une valeur de 0,80 signifie que les décisions sont exactes ou constantes pour 80 % des candidats.

Tableau 6 – Estimation de fidélité, erreur type de mesure et indices de constance et d'exactitude des décisions pour chacune des épreuves de 2017

	MARS		SEPTEMBRE		
	Épr. n° 1	Épr. n° 2	Épr. n° 3	Épr. n° 4	Épr. n° 5
Estimation de fidélité	0,79	0,68	0,73	0,74	0,73
ETM (notes sur l'échelle des scores communiquées)	3,42	2,74	4,08	3,32	4,01
Constance des décisions	0,91	0,93	0,82	0,94	0,86
Faux positifs	0,04	0,04	0,10	0,03	0,07
Faux négatifs	0,04	0,03	0,08	0,03	0,07
Exactitude des décisions	0,94	0,95	0,87	0,96	0,90
Faux positifs	0,02	0,02	0,05	0,01	0,03
Faux négatifs	0,04	0,03	0,07	0,03	0,07

Statistiques pour les stations d'ECOS

Le tableau 7 donne des statistiques sommaires pour chacune des stations d'ECOS des épreuves de 2017. Ces statistiques comprennent le pourcentage de données manquantes, les proportions de bonnes réponses (les valeurs P), l'écart-type des notes et la corrélation totale des stations (CTS).

Les valeurs P représentent les proportions moyennes de bonnes réponses obtenues par les candidats. En général, les valeurs P, qui vont de 0 à 1, sont un indicateur du niveau de difficulté des stations. Une valeur P faible (< 0,20) est le signe d'une station difficile, et une valeur P élevée (> 0,90) est le signe d'une station facile. Les valeurs P dépendent de la cohorte de

candidats. Toute comparaison des valeurs P entre divers échantillons de candidats ne tient pas compte des différences potentielles au niveau des compétences globales. Il ne faut donc pas accorder trop d'importance à l'interprétation des valeurs P ni les utiliser comme indicateur exclusif du niveau de difficulté des stations. Elles montrent plutôt de manière générale le niveau variable de difficulté des stations dans une version donnée d'une épreuve.

Les écarts-types indiquent la variabilité des notes obtenues aux diverses stations. Les corrélations totales des stations (CTS) sont des indicateurs des différences entre les candidats dont le rendement est élevé et ceux dont le rendement est faible, pour une station donnée. Une valeur de CTS faible ou négative ($< 0,15$) signifie que les candidats dont le niveau de rendement est élevé ne réussissent pas aussi bien qu'ils le devraient dans une station d'ECOS donnée et que les candidats dont le niveau de rendement est faible réussissent mieux que leurs résultats laisseraient prédire. Cette information, combinée avec les valeurs P, peut être le signe de stations qui devraient être éliminées du processus de notation, puisqu'elles ne permettent pas une bonne discrimination entre les candidats de niveau élevé et de niveau faible. Une valeur de CTS modérée à élevée ($> 0,30$) indique que les candidats dont le niveau est élevé réussissent bien dans une station d'ECOS donnée en comparaison aux candidats dont le niveau est faible. Les stations dont la valeur de CTS est inférieure à 0,15 ou négative sont signalées pour qu'elles soient passées en revue. Les stations dont le contenu est jugé acceptable sont retenues.

Tableau 7 – Statistiques globales sur les stations d'ECOS pour chacune des épreuves de 2017

MARCH				
Épreuve n° 1				
Station	Données manquantes (%)	Valeur P	Écart-type	CTS
2	0,78	0,53	0,16	0,54
3	0,00	0,61	0,15	0,46
4	0,06	0,66	0,15	0,42
5	0,15	0,61	0,15	0,52
6	0,04	0,58	0,16	0,42
7	0,04	0,65	0,16	0,48
9	0,00	0,59	0,19	0,35
10	0,13	0,59	0,18	0,49
11	0,08	0,59	0,14	0,51
12	0,19	0,66	0,14	0,46
Moyenne	0,18	0,61	0,16	0,47

SEPTEMBRE								
Station	Épreuve n° 2				Épreuve n° 3			
	Données manquantes (%)	Valeur P	Écart-type	CTS	Données manquantes (%)	Valeur P	Écart-type	CTS
2	0,06	0,44	0,17	0,46	0,05	0,61	0,16	0,36
3	0,07	0,55	0,15	0,48	0,19	0,58	0,13	0,44
4	0,00	0,54	0,16	0,34	0,00	0,56	0,16	0,34
5	0,35	0,58	0,14	0,34	0,05	0,58	0,15	0,39
6	0,30	0,61	0,15	0,36	0,08	0,56	0,13	0,42
7	0,00	0,58	0,11	0,46	0,14	0,53	0,16	0,30
9	0,00	0,63	0,14	0,27	0,00	0,57	0,19	0,38
10	0,26	0,60	0,16	0,21	0,23	0,53	0,15	0,37
11	0,19	0,59	0,13	0,21	0,28	0,54	0,13	0,47
12	0,06	0,69	0,11	0,36	0,00	0,55	0,16	0,49
Moyenne	0,18	0,58	0,14	0,35	0,14	0,56	0,15	0,40

Tableau 7 (suite) – Statistiques globales sur les stations d'ECOS pour chacune des épreuves de 2017

Station	SEPTEMBRE							
	Épreuve n° 4				Épreuve n° 5			
	Données manquantes (%)	Valeur P	Écart-type	CTS	Données manquantes (%)	Valeur P	Écart-type	CTS
2	0,07	0,57	0,13	0,46	0,00	0,59	0,14	0,46
3	0,11	0,59	0,14	0,41	0,00	0,59	0,14	0,27
4	0,10	0,54	0,17	0,45	0,05	0,62	0,12	0,36
5	0,88	0,64	0,12	0,45	0,06	0,66	0,13	0,38
6	0,00	0,64	0,13	0,37	0,00	0,58	0,16	0,25
7	0,03	0,63	0,13	0,41	0,00	0,65	0,13	0,45
9	0,00	0,52	0,15	0,32	0,00	0,56	0,18	0,44
10	0,00	0,62	0,14	0,35	0,00	0,65	0,12	0,53
11	0,07	0,62	0,16	0,34	0,10	0,61	0,12	0,47
12	0,13	0,70	0,17	0,44	0,10	0,68	0,15	0,41
Moyenne	0,11	0,61	0,14	0,40	0,08	0,62	0,14	0,40

Les valeurs de CTS de trois stations de l'épreuve no 2 et deux stations de l'épreuve 5 étaient inférieures à 0,30 et ont été signalées pour être passées en revue. Après avoir été examinées par des experts du contenu médical, elles ont été jugées acceptables et ont été incluses dans le calcul des scores.

Profil des résultats par compétence

Un profil des résultats par compétence est présenté sous forme graphique à chaque candidat dans le Rapport d'information supplémentaire. Des exemples du Relevé de résultats et du Rapport d'information supplémentaire sont disponibles à l'annexe A et B, respectivement. Le graphique du Rapport d'information supplémentaire montre le score par compétence obtenu par le candidat pour chacune des neuf compétences (capacité de communiquer, interprétation des données, diagnostic, prise d'anamnèse, analyses, maîtrise

de la langue, organisation, examen physique, et prise en charge thérapeutique), ainsi que l'erreur type de mesure correspondante. Le calcul des scores par compétence pour chaque candidat est expliqué dans la section Calcul des scores par compétence du chapitre de la Notation de l'examen de la CNE de ce rapport.

La présente section donne le profil des résultats par compétence pour chacune des cinq épreuves de 2017. La fourchette des résultats par compétence est représentée sous forme graphique aux figures 3 à 7. Pour chaque compétence, le rectangle représente la fourchette des scores obtenus par 50 pour cent des candidats. La ligne verticale à l'intérieur de chaque rectangle correspond à la note médiane ou au 50^e centile des scores par compétence. Les 25 pour cent supérieurs et inférieurs restants des scores par compétence sont représentés par des lignes horizontales, respectivement à la droite et à la gauche du rectangle. Le losange indique la note moyenne par compétence. Sous chaque figure, la légende donne la signification des abréviations utilisées pour chaque compétence.

En résumé, les figures 3 à 7 illustrent les scores par compétence pour chacune des versions d'épreuves et pour l'ensemble des candidats. Ces graphiques montrent que certains scores par compétence sont plus variables (par exemple, que l'étendue des résultats est très grande, avec des scores très élevés et d'autres très faibles). De plus, les graphiques illustrent les forces et faiblesses relatives pour chacune des compétences évaluées dans chacune des épreuves. Il est à noter que les profils des résultats par compétence ne peuvent être comparés d'une épreuve à l'autre, car ils sont dressés individuellement pour chaque épreuve et ne sont pas liés statistiquement à l'ensemble des épreuves comme le sont les scores d'examen totaux.

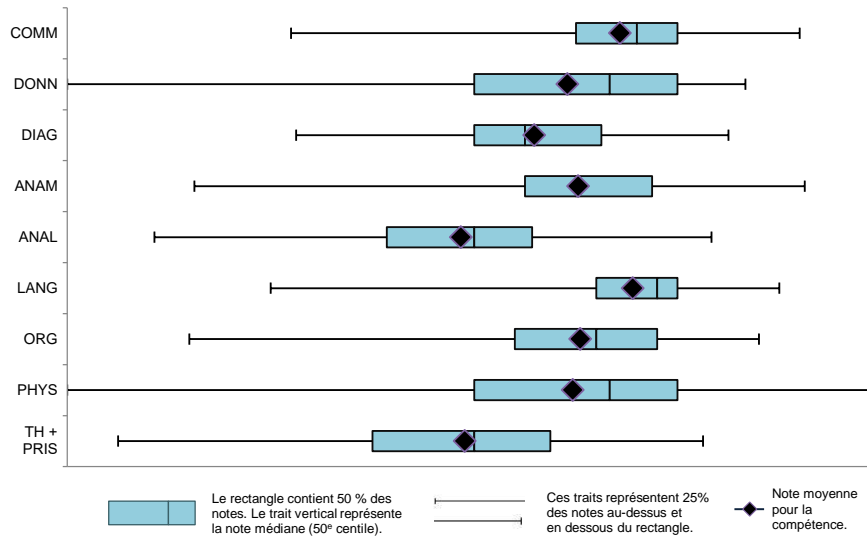


Figure 3 – Profil des résultats par compétence à l'épreuve n° 1 de 2017 (tous les candidats)

Légende : COMM = capacité de communiquer, DONN = interprétation des données, DIAG = diagnostic, ANAM = prise d'anamnèse, ANAL = analyses, LANG = maîtrise de la langue, ORG = organisation, PHYS = examen physique, TH + PRIS = thérapeutique et prise en charge

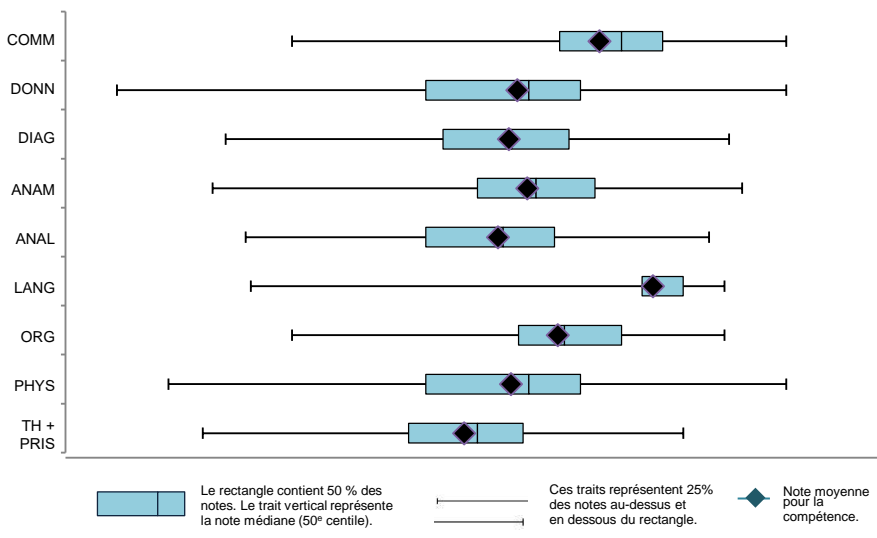


Figure 4 – Profil des résultats par compétence à l'épreuve n° 2 de 2017 (tous les candidats)

Légende : COMM = capacité de communiquer, DONN = interprétation des données, DIAG = diagnostic, ANAM = prise d'anamnèse, ANAL = analyses, LANG = maîtrise de la langue, ORG = organisation, PHYS = examen physique, TH + PRIS = thérapeutique et prise en charge

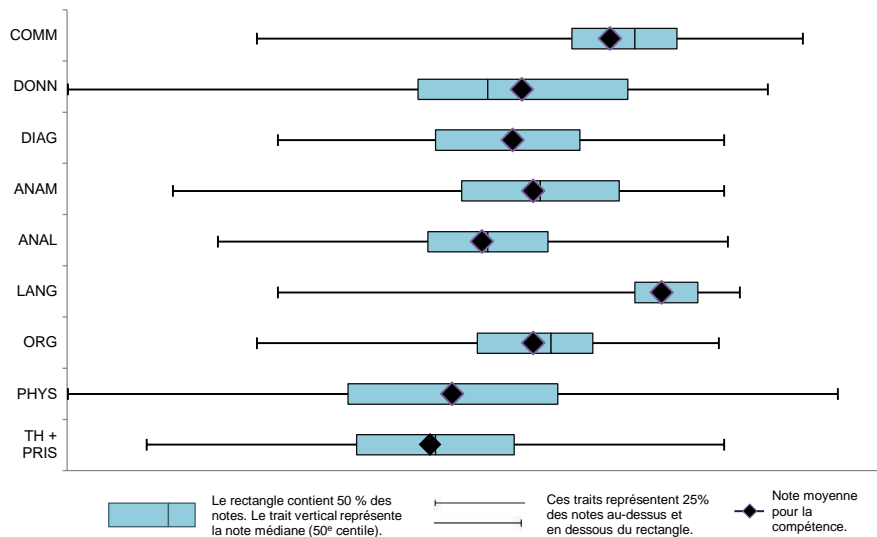


Figure 5 – Profil des résultats par compétence à l'épreuve n° 3 de 2017 (tous les candidats)

Légende : COMM = capacité de communiquer, DONN = interprétation des données, DIAG = diagnostic, ANAM = prise d'anamnèse, ANAL = analyses, LANG = maîtrise de la langue, ORG = organisation, PHYS = examen physique, TH + PRIS = thérapeutique et prise en charge

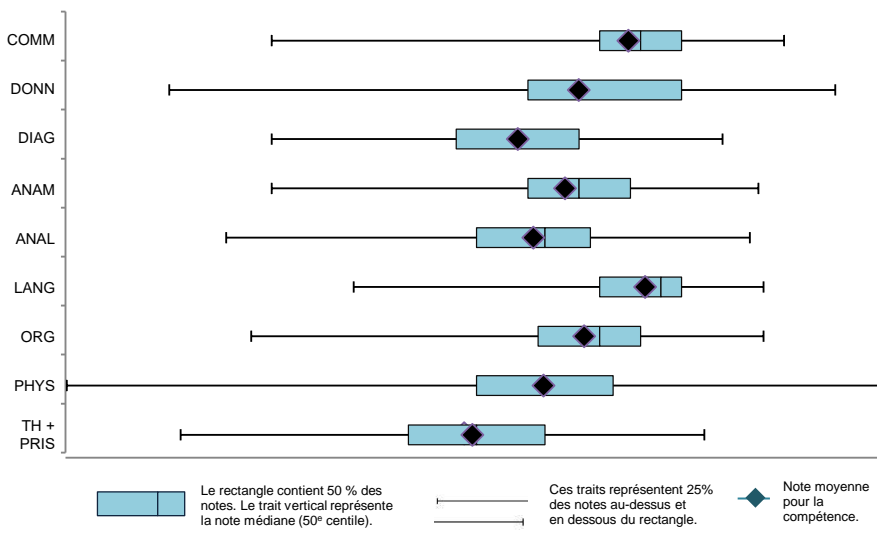


Figure 6 – Profil des résultats par compétence à l'épreuve n° 4 de 2017 (tous les candidats)

Légende : COMM = capacité de communiquer, DONN = interprétation des données, DIAG = diagnostic, ANAM = prise d'anamnèse, ANAL = analyses, LANG = maîtrise de la langue, ORG = organisation, PHYS = examen physique, TH + PRIS = thérapeutique et prise en charge

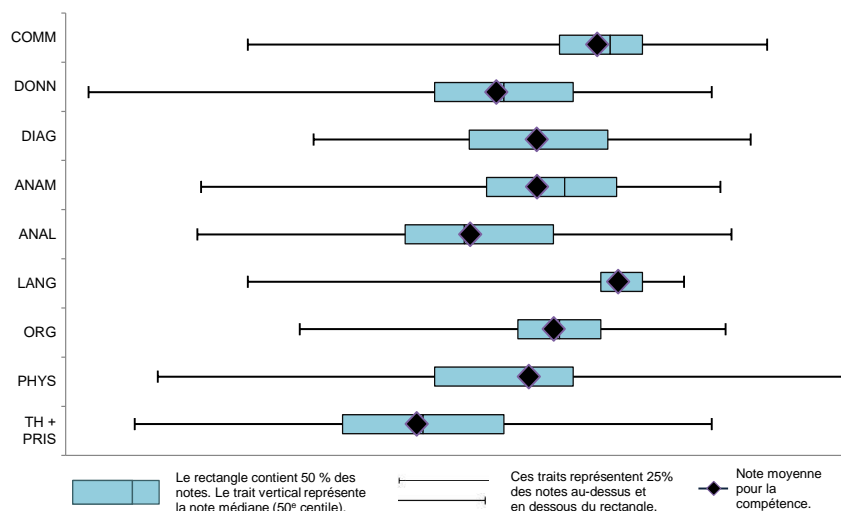


Figure 7 – Profil des résultats par compétence à l'épreuve n° 5 de 2017 (tous les candidats)

Légende : COMM = capacité de communiquer, DONN = interprétation des données, DIAG = diagnostic, ANAM = prise d'anamnèse, ANAL = analyses, LANG = maîtrise de la langue, ORG = organisation, PHYS = examen physique, TH + PRIS = thérapeutique et prise en charge

Comparaisons historiques

Le tableau 8 donne, à des fins de comparaison, les taux de réussite des candidats aux examens de 2013 à 2017.

Tableau 8 – Taux de réussite des candidats de 2013 à 2017

Année	Session	TOTALS		Candidats pour la 1 ^{ère} fois		Candidats pour la 2 ^e fois ou plus	
		Nbre	Taux de réussite	Nbre	Taux de réussite	Nbre	Taux de réussite
2017	Total	1506	86,79 %	1430	87,76 %	76	68,42 %
	Mars	330	87,88 %	292	89,04 %	38	78,95 %
	Sept.	1176	86,48 %	1138	87,43 %	38	57,89 %
2016	Total	1547	92,63 %	1447	93,99 %	100	73,00 %
	Mars	380	91,84 %	327	94,50 %	53	75,47 %
	Sept.	1167	92,89 %	1120	93,84 %	47	70,21 %
2015	Total	1487	91,19 %	1392	91,95 %	95	80,00 %
	Mars	335	84,48 %	295	84,07 %	40	87,50 %
	Sept.	1152	93,14 %	1097	94,07 %	55	74,55 %
2014	Total	2106	92,64 %	1680	93,00 %	426	92,00 %
	Mars	420	83,81 %	296	82,77 %	124	86,29 %
	Sept.	1686	94,84 %	1384	95,09 %	302	93,71 %
2013	Total	1221	77,58 %	986	76,00 %	235	85,00 %
	Mars	547	70,93 %	439	68,79 %	108	79,63 %
	Sept.	674	83,09 %	547	81,54 %	127	89,76 %

5. RÉFÉRENCES BIBLIOGRAPHIQUES

- Bahry *et al.*, 2012 – Bahry, L., R.K. Hambleton, A. Gotzmann, A.F. De Champlain et M. Roy. *National Assessment Collaboration Standard Setting Study Report*, rapport technique pour le comité consultatif technique et la Collaboration nationale en matière d'évaluation, Ottawa, CMC.
- CCPSSRH, 2004 – Comité consultatif sur la prestation des soins de santé et les ressources humaines. *Rapport du Groupe de travail sur le permis d'exercice des diplômés internationaux en médecine*, février 2004.
- Cizek et Bunch, 2007 – Cizek, G.J., et M.B. Bunch. *Standard setting: A guide to establishing and evaluating performance standards on tests*, Thousand Oaks, CA, Sage Publications Inc.
- Crocker et Algina, 1986 – Crocker, L., et J. Algina. *Introduction to classical and modern test theory*. Orlando, FL, Harcourt Brace Jovanovich College Publishers.
- Federal/Provincial/Territorial Advisory Committee on Health Delivery and Human Resources. *Report from the Canadian Task Force on Licensure of International Medical Graduates* (February 2004).
- Haertel, 2006 – Haertel, E.H. « Reliability », dans R.L. Brennan (réd.), *Educational measurement*, 4^e édition, Westport, CT, Praeger, p. 65-110.
- Hambleton et De Champlain, 2013 – Hambleton, R.K., et A.F. De Champlain. *Recommending a passing score on the National Assessment Collaboration (NAC) examination: Procedures and results*, rapport technique pour le comité d'épreuve de la Collaboration nationale en matière d'évaluation, Ottawa, CMC.
- Kane, 2006 – Kane, M.T. « Validation », dans R. L. Brennan (réd.), *Educational measurement*, 4^e édition, Westport, CT, Praeger, p. 17-64.
- Kane, 2013 – Kane, M.T. « Validating the interpretations and uses of test scores », *Journal of Educational Measurement*, vol. 50, n° 1, p. 1-73.
- Kolen et Brennan, 2004 – Kolen, M.J., et R.L. Brennan. *Test equating, scaling, and linking: Methods and practice*, 2^e édition, New York, NY, Springer Science Business Meeting.
- Livingston et Lewis, 1995 – Livingston, S.A., et C. Lewis. « Estimating the consistency and accuracy of classifications based on test scores », *Journal of Educational Measurement*, vol. 32, n° 2, p. 179-197. doi: 10.1111/j.1745-3984.1995.tb00462.x

ANNEXE A : RELEVÉ DE RÉSULTATS POUR L'EXAMEN DE LA CNE



L'EXAMEN DE LA COLLABORATION NATIONALE EN MATIÈRE D'ÉVALUATION RELEVÉ DE RÉSULTATS

Le 24 février 2016

Nom du candidat :	xxxxxxxx, xxxxx	Votre résultat final :	Réussite
Matricule du CMC :	xxxxxxxxxx	Votre note globale :	75
Date de l'examen :	2016-01-19	Note de passage exigée :	65

Je vous écris au nom du Comité d'épreuve de la Collaboration nationale en matière d'évaluation afin de vous informer de votre statut quant à vos résultats à l'examen d'évaluation CNE ayant eu lieu à la date mentionnée ci-dessus.

Votre note globale, qui représente votre rendement global à cet examen, est rapportée sur une échelle de notes allant de 0 à 100. Votre résultat final (c'est-à-dire réussite / échec) est basé sur votre note globale en lien avec la note relative nécessaire pour réussir l'examen. Des renseignements supplémentaires, incluant la moyenne et l'écart-type, sont disponibles sur la page Notation du site mcc.ca.

De l'information supplémentaire à propos de votre performance lors de l'examen vous sera transmise dans un document séparé sur votre compte inscriptionmed.ca.

Veillez accepter mes meilleurs vœux de succès pour l'avenir.

M. Ian Bowmer, MD CM, FRCPC
Directeur général et registraire
Le Conseil médical du Canada

ANNEXE B : RAPPORT D'INFORMATION SUPPLÉMENTAIRE POUR L'EXAMEN DE LA CNE



RAPPORT D'INFORMATION SUPPLÉMENTAIRE

Nom du candidat : xxxxxxxx, xxxxxx

Matricule du CMC : xxxxxxxxxx

Votre note globale : 75

Examen : CNE

Date de l'examen : 2016-01-19

L'objectif du présent rapport est de vous fournir des renseignements supplémentaires sur vos forces et faiblesses relatives à votre rendement dans les différents domaines qui ont été évalués dans le cadre du formulaire de l'examen de la Collaboration nationale en matière d'évaluation (CNE) qui vous a été administré.

La Figure 1 présente votre rendement tel que mesuré aux stations relativement à chacun des neuf domaines suivants : capacité de communiquer, interprétation des données, diagnostic, prise de l'anamnèse, analyses, maîtrise de la langue, organisation, examen physique et connaissances thérapeutiques (TPx) et prise en charge du patient. Chaque domaine est évalué un certain nombre de fois, certains l'étant à toutes les stations, et d'autres, à un nombre restreint de stations.

Afin de vous aider à mieux comprendre votre rendement, vos notes partielles propres à chaque domaine sont indiquées et assorties des notes moyennes des candidats de la même session d'examen qui passaient l'examen de la CNE pour la première fois et qui l'ont réussi. L'erreur type de mesure (ETM) associée à chacune de vos notes partielles représente l'écart prévu de votre note partielle si vous deviez passer l'examen avec un ensemble différent de stations couvrant les mêmes domaines ou des domaines similaires. Des notes partielles légèrement différentes ou des chevauchements entre les ETM révèlent que le rendement dans ces domaines était relativement semblable. De même, un chevauchement entre l'ETM de la note partielle dans un domaine et la note moyenne des candidats qui passaient l'examen pour la première fois et qui l'ont réussi, dans un domaine donné, signifie que le rendement était semblable au rendement moyen.

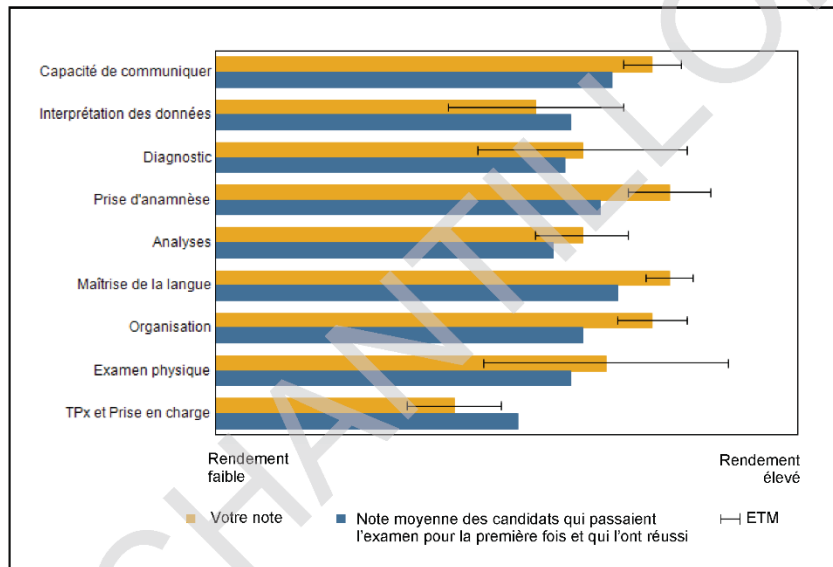
Date du rapport : 2016-02-24

1/2

Il convient de signaler que les notes partielles sont fondées sur un moins grand nombre de données que la note globale et qu'elles n'affichent pas le même niveau de précision que la note globale. Si vous avez échoué à l'examen et que vous voulez le reprendre, la préparation pour tous les domaines est importante; sinon, vous pourriez améliorer certaines notes partielles, mais en réduire d'autres par inadvertance. Il convient aussi de signaler que vos notes partielles ne peuvent pas être comparées directement à votre note globale puisqu'elles sont calculées différemment et que, par conséquent, la note de passage ne peut pas être reproduite dans la Figure 1.

Pour de plus amples renseignements, veuillez consulter la [page Web des examens de la CNE](#).

* Figure 1 : Profil des notes à l'examen de la CNE



Date du rapport : 2016-02-24

2/2

xxxxxxxx, xxxxx / xxxxxxxxxxx