

Etude d'impact de séances d'entraînement à des exercices de dossiers cliniques simulés sur les performances d'étudiants à un examen national

Study of the Impact of training sessions of simulated clinical charts exercises on the performance of students during a national examination.

Pierre POTTIER¹, Bernard PLANCHON¹, Jacques Henri BARRIER¹, Christelle VOLTEAU², Alain MOUZARD³

Résumé **Contexte** : Des Séances hebdomadaires d'entraînement au raisonnement médical (SERAM) basées sur la résolution de problèmes cliniques ont été mises en place pendant les trois années du cursus clinique prégradué de notre faculté (4^e à 6^e années d'études). **But/sujets/matériel** : Nous avons comparé les étudiants ayant suivi régulièrement ces séances d'entraînement (groupe A) aux étudiants non assidus (groupe B). La comparaison a été réalisée sur 4 critères : 1) le score moyen à une simulation d'examen classant et 2) aux examens de fin d'année en 4^e année, 3) le rang de classement moyen aux examens de fin de 6^e année et 4) aux épreuves classantes nationales (ECN). **Résultats** : Le niveau entre les 2 groupes était identique lors du concours de sélection de première année. On note une amélioration progressive et très précoce (dès la fin de la 4^e année) des performances des étudiants du groupe A aux différents examens puisqu'ils obtiennent une meilleure note sur 2 dossiers à l'examen blanc de fin de 4^e année, 5 dossiers à l'examen de fin d'année en 4^e année et une amélioration du rang de classement de 500 places aux ECN. **Conclusion** : Même s'il n'est pas exclu que les étudiants assidus soient a priori les plus motivés et les plus laborieux, ces SERAM semblent améliorer les performances des étudiants à un examen classant national. Il reste à prouver que ce type d'enseignement basé sur des scripts construits par des cliniciens expérimentés décrivant des situations cliniques complexes multiples et variées permet d'augmenter les compétences d'exercice de nos futurs collègues.

Mots clé Dossiers cliniques simulés ; raisonnement clinique évaluation formative ; épreuves classantes nationales.

Abstract **Background**: In our university, 4th to 6th year medical students exercise once a week their medical reasoning by attending to courses based on clinical problem resolution derived from real clinical situations. **Aim/subject/material**: We compared students who attended to these training courses on a regular basis (group A) to students who did not (group B). The comparison took into account 4 criteria: the average mark at 1) the mock exam, 2) the final exam of fourth-year students, 3) the final exam of sixth-year students and 4) the national ranking test. **Results**: After the first year exam, the level of performance between both groups was identical. However, we have noticed an early and progressive improvement of students in the group A at different exams: they obtained a better mark in two subjects at the mock exam and in five subjects at the sixth-year final exams. They also have gained 500 places in the national ranking. **Conclusion**: Even if the students who attended to these courses are considered to be the most motivated and hard-working learners, we can nonetheless suppose that these courses seemed to have improved their performances. It remains to be proven that this teaching method based on scripts elaborated by experienced clinicians, which describe complex cross disciplinary clinical situations, is a way to increase the medical performance of our future clinicians.

Key-words Clinical reasoning; patients' management problem; formative evaluation; national ranking exam.

Pédagogie Médicale 2006;7:213-27

Introduction

L'activité de résolution de problèmes de santé en médecine, couramment appelée raisonnement clinique, est l'une des composantes essentielle de la compétence clinique attendue des futurs médecins diplômés. Grâce à de nombreux travaux de recherche, les processus en jeu sont aujourd'hui beaucoup mieux connus et des pistes pour organiser son enseignement et son apprentissage ont été identifiées^{1, 2, 3}.

La problématique du raisonnement clinique

De nombreux concepts tentent actuellement de « modéliser » le raisonnement clinique^{1, 2, 3, 4, 5} sans qu'aucun ne permette de rendre compte de manière satisfaisante de l'ensemble des processus mis en jeu lors de la construction d'un diagnostic et de la prise de décision, ce d'autant que ces processus semblent varier selon l'expertise, les connaissances théoriques de base et la complexité des problèmes cliniques à résoudre.

Même si cette distinction est en partie la conséquence de limites d'ordre méthodologique¹, il reste classique de considérer que le raisonnement clinique comporte deux étapes successives : la résolution diagnostique et la prise de décision.

La résolution diagnostique

Plusieurs approches d'ailleurs non antinomiques ont été proposées pour tenter d'explicitier les procédures cognitives élaborées au cours de la résolution d'un problème clinique. Il apparaît que ces méthodes de résolution de problèmes varient selon la difficulté du cas ressenti par le clinicien.

Dans le cas de problèmes identifiés comme simples (souvent parce qu'ils relèvent très spécifiquement du champ de compétence du médecin consulté), le diagnostic est établi le plus souvent par catégorisation ou reconnaissance de cas ou encore analyse propositionnelle^{6, 7, 8, 9}. Le médecin confronte les données du patient à un répertoire ordonné de cas déjà rencontrés ou d'exemples particulièrement typiques. Il peut également utiliser une méthode plus abstraite dans laquelle il compare le patient à des prototypes

construits mentalement grâce à un processus d'abstraction élaboré à partir de « vrais » patients¹⁰.

Dans le cas de problèmes identifiés comme complexes, le clinicien met en jeu des processus différents qui peuvent partir des symptômes pour aller au diagnostic (processus inductif) ou à l'inverse partir de l'hypothèse pour revenir à la collecte orientée d'information (processus hypothético-déductif). Le premier processus sollicite des représentations cognitives complexes telles que les réseaux sémantiques¹¹ reliés entre eux par des nœuds décisionnels permettant de trier les hypothèses. Processus inverse, la démarche hypothético-déductive se fait en deux étapes, la génération précoce d'hypothèses précédant l'évaluation de leur probabilité¹². L'évaluation des hypothèses émises et, *in fine*, leur hiérarchisation s'appuient sur une collecte orientée de nouvelles données visant soit à les accréditer, soit à les discréditer.

La prise de décision

Le processus de prise de décision commence par la sélection d'un diagnostic parmi plusieurs possibles. Certaines perspectives normatives, telles que la théorie bayésienne¹³, stipulent que sélectionner l'hypothèse la plus probable devrait conduire à comparer des probabilités pré test (prévalence de la maladie) et des probabilités post test (après l'examen clinique et paraclinique). Cette action implique d'être capable d'interpréter les signes cliniques ou les tests paracliniques, c'est-à-dire de maîtriser les notions de sensibilité, spécificité et valeurs prédictives.

La décision médicale ne concerne pas que les aspects diagnostiques. A l'étape thérapeutique, il faut dans un premier temps faire un choix entre initier un traitement, le différer ou encore s'en abstenir ; s'il est décidé d'introduire un traitement, il faut savoir sélectionner le meilleur dans l'indication donnée, compte tenu notamment du terrain du patient c'est-à-dire savoir établir une balance bénéfice-risque. Chercher à expliciter les éléments de la décision médicale c'est donc explorer le domaine de la médecine basée sur les preuves.

Cependant, la décision médicale aborde aussi d'autres aspects moins « biomédicaux » que sont la sélection des informations données au patient et l'utilisation d'une communication adaptée à chaque patient. Ceci nécessite l'acquisition d'une expertise dans le domaine de la relation

1- Service de médecine interne - CHU de Nantes - Place Alexis Ricordeau - 44093 Nantes Cedex - Unité de formation et de recherche de médecine de Nantes.

2- Pôle d'information médicale d'évaluation et de santé publique (P.I.M.E.S.P) - Hôpital Saint Jacques - 44093 Nantes Cedex 1.

3- Service de Pédiatrie - Hôpital Mère et Enfant - 44093 Nantes Cedex 1 - Unité de Formation et de Recherche de Médecine de Nantes.

Correspondance : Pierre Pottier - Service de médecine interne - Place Alexis Ricordeau - 44093 Nantes Cedex - Téléphone : 02 40 08 33 52

Mailto:pierre.pottier@univ-nantes.fr

médecin-patient avec tous ses aspects psycho-socio-professionnels et ethnologiques, articulée avec des compétences à prendre en compte les dimensions éthiques de l'action médicale.

La problématique de l'enseignement, de l'apprentissage et de l'évaluation du raisonnement clinique

Ainsi, même s'il persiste encore plusieurs questions non résolues, de nombreuses données sont désormais disponibles concernant les activités de résolution de problèmes que développent les professionnels de santé et les processus mentaux qu'ils mettent en œuvre pour ce faire. Un autre défi, pour les enseignants, est de concevoir et de mettre en œuvre des activités d'apprentissage, d'enseignement et d'évaluation qui créent les conditions les plus favorables à la construction progressive des compétences appropriées par les étudiants.

Les dispositifs de formation élaborés dans cette intention s'efforcent d'être en cohérence avec les acquis récents des sciences de l'apprentissage, qui conduisent à recommander d'exposer les étudiants à des tâches authentiques, complexes, complètes et significatives^{1, 2, 3}, qui leur permettent d'expérimenter et de développer à la fois des stratégies analytiques et non analytiques de raisonnement clinique. Ces activités pédagogiques peuvent être formelles ou informelles, organisées en contexte réel de stage ou sous forme d'activités simulées, telles que le permettent par exemple les séances formalisées d'apprentissage du raisonnement clinique (ARC)¹⁴ ou encore des études de dossiers cliniques. Dans tous les cas, il est recommandé de permettre aux étudiants de développer simultanément les stratégies de raisonnement et les connaissances spécifiques. A cet égard, la fourniture d'une rétroaction explicite de la part des enseignants vers les étudiants est tenue pour être un élément déterminant^{1, 2, 3}.

Concernant l'évaluation du raisonnement clinique, de nombreux outils ont été développés et étudiés, ce qui a permis d'en établir les forces et les faiblesses¹⁵. Pour tenter de dépasser certaines limites et, par exemple, afin de mieux prendre en compte la dimension d'incertitude inhérente au raisonnement clinique, des formats spécifiques d'évaluation tels que le test de concordance de scripts ont été élaborés, qui visent à explorer des tâches cognitives plus pertinentes¹⁶. D'une façon générale, on considère que la validité et la fidélité de la démarche évaluative requièrent que l'on utilise plusieurs outils et formats d'évaluation et qu'on y soumette les étudiants

de façon récurrente, en explorant un échantillon suffisamment varié de situations cliniques¹⁵.

Le contexte académique réglementaire français et le contexte institutionnel de la faculté de médecine de Nantes

Le contexte académique réglementaire français

En France, des dispositions réglementaires^{17, 18}, opposables à toutes les facultés de médecine, ont récemment à la fois modifié les orientations et les finalités des trois années du deuxième cycle des études médicales (DCEM) et fixé les modalités des épreuves classantes nationales (ECN), auxquelles doivent se présenter tous les étudiants pour accéder au troisième cycle. Les finalités assignées au DCEM visent explicitement « [...] l'acquisition des compétences cliniques et thérapeutiques [...] », en prévoyant notamment que « [...] en fin de deuxième cycle, tous les étudiants doivent avoir assimilé [...] » : les pathologies les plus fréquentes, leurs procédures diagnostiques, leurs thérapeutiques et leurs préventions ; la démarche médicale en fonction de la prévalence, de la gravité et des possibilités thérapeutiques ; la gestion des urgences les plus fréquentes ; la maîtrise des outils de la relation et de la communication [et qu'] il est tout aussi essentiel que les étudiants soient aptes à informer les patients et leurs familles en termes simples et compréhensibles, pour mieux les associer aux décisions qui les concernent ».

Pour leur part, les ECN ne comportent qu'un seul format d'évaluation, consistant en l'analyse de neuf dossiers cliniques explorés par des questions à réponse ouverte et courtes. Les conséquences du caractère univoque de ce format d'évaluation seront discutées plus loin. Il convient simplement de rappeler que les ECN visent uniquement à classer tous les étudiants français à l'issue de leur second cycle et qu'elles n'ont pas explicitement de fonction certificative, celle-ci étant dévolue aux épreuves organisées dans chacune des facultés tout au long du deuxième cycle¹⁷, et notamment à la fin de celui-ci dans le cadre du certificat de synthèse clinique et thérapeutique.

Le contexte institutionnel de la faculté de médecine de Nantes

Depuis la parution de l'arrêté ministériel du 10 octobre 2000, la faculté de médecine de Nantes a institué des séances d'entraînement au raisonnement médical (SERAM) basées sur la résolution de problèmes^{9, 19} dans un double objectif : préparer les étudiants aux ECN et approfondir grâce à des questions ciblées certains

aspects du raisonnement clinique évoqués plus haut. Nous avons comparé les performances aux examens d'étudiants ayant suivi régulièrement ces séances d'entraînement à celles d'étudiants ayant choisi un autre mode de préparation (préparation individuelle ou dans le cadre de séances moins formelles organisées sous la supervision d'internes).

Objectifs de l'étude

Nous avons testé l'hypothèse suivante : « plus souvent les étudiants participent aux SERAM lors de la première année (en quatrième année des études médicales), plus leur performance aux ECN est élevée ». Si tel est le cas, la comparaison des résultats aux divers examens facultaires et aux ECN des étudiants les plus assidus aux SERAM par rapport aux étudiants les moins assidus devrait confirmer le bénéfice de la méthode.

Méthodes

Description de l'outil pédagogique étudié

L'arrêté ministériel du 10 octobre 2000¹⁷ établit de nouveaux objectifs d'apprentissage, regroupés pour certains en modules transversaux interdisciplinaires. Le nouveau programme a pour objectif une formation plus transversale que le concours précédent, se voulant préparatoire aux diplômes d'études spécialisées dans le cadre desquels s'intègre désormais la formation des futurs médecins généralistes.

Les recommandations pour la rédaction des dossiers cliniques des ECN émises par le Conseil scientifique du centre national du concours de l'internat (CNCI) (cf. *Annexe 1*) affichent explicitement une volonté d'explorer certaines dimensions du raisonnement clinique.

Les questions à choix multiples ont été abandonnées, bien que, en dépit de certaines limites, elles ne soient pas du tout disqualifiées pour explorer le raisonnement clinique¹⁵. Les neuf dossiers des ECN sont théoriquement construits dans un souci de transversalité interdisciplinaire, visant à simuler des situations complexes pour tester les capacités de l'étudiant à hiérarchiser et argumenter ses hypothèses et à sélectionner les examens complémentaires pertinents en prenant en compte le patient dans sa globalité. Cependant, un certain nombre de dimensions du raisonnement médical ne paraissent pas pouvoir être évaluées par de tels dossiers, *a fortiori* dans le contexte d'une évaluation normative. Ainsi, ils sont inadaptés pour apprécier les aptitudes à développer une collecte orientée de données, à gérer l'incertitude, à éduquer et à informer le patient ou sa famille, dans la mesure où ils fournissent à

l'étudiant un nombre fini de données et qu'aucune donnée complémentaire ne peut être intentionnellement recherchée.

Pour mieux préparer la première promotion d'étudiants soumise à cette évaluation, des SERAM ont été proposées aux étudiants de manière facultative et hebdomadaire tout au long du second cycle, couvrant les 345 thèmes du programme officiel des ECN. Les thèmes de chaque séance (cf. *Annexe 2*), conformes au programme de l'année en cours, étaient affichés à l'avance pour permettre un travail personnel préalable. Les étudiants disposaient de deux heures pour prendre connaissance de deux à quatre problèmes cliniques, élaborés selon les recommandations du CNCI et pour rédiger leurs réponses. Aussi souvent que possible, les questions ont été optimisées, comme le proposent Charlin et *al.*¹⁵, pour l'évaluation du raisonnement clinique en exigeant par des réponses courtes et ciblées de justifier une attitude ou une décision d'ordre médical (diagnostique, thérapeutique ou sélection d'un examen complémentaire). Une rétroaction était assurée par l'enseignant pendant les deux heures suivantes à partir d'une grille détaillée et critériée, qui était commentée pendant la correction et remise aux étudiants en fin de séance. Chaque fois que possible, la capacité de l'étudiant à gérer l'incertitude était testée grâce à des grilles suffisamment souples pour permettre d'accorder le même crédit à des réponses différentes à partir du moment où elles étaient parfaitement argumentées (cf. *Annexe 3*). Le référentiel d'apprentissage était indiqué à l'avance pour permettre à l'étudiant d'actualiser ses connaissances déclaratives considérées comme nécessaires à la résolution des exercices proposés. Le programme des SERAM a été construit de manière à ce que de nombreux thèmes puissent être abordés plusieurs fois par des problèmes cliniques différents, ceci pour faciliter un travail de contextualisation et re-contextualisations itératives de la part de l'étudiant. Pour une meilleure assiduité à ces séances (le mode d'évaluation conditionnant le mode d'apprentissage), la faculté de Nantes a modifié totalement ses méthodes docimologiques depuis 2001. Ainsi, les étudiants sont actuellement évalués, dans les conditions des ECN, sur sept, huit et neuf dossiers respectivement en fin de deuxième (DCEM2), troisième (DCEM3) et quatrième (DCEM4) années du deuxième cycle.

Définition de l'assiduité des étudiants - Constitution des groupes

Assiduité

A l'issue de chacune des séances non obligatoires, les étudiants devaient obligatoirement remplir et remettre à

Etude d'impact de séances d'entraînement...

l'enseignant une fiche d'auto-évaluation (résultats non rapportés) sur laquelle leurs scores aux dossiers devaient être reportés après auto-correction d'après les grilles fournies par l'enseignant. Cette fiche d'auto-évaluation permettait également de contrôler leur assiduité.

Constitution des groupes

Deux groupes ont été constitués, d'après le nombre de fiches d'auto-évaluation rendues lors de la première année (soit en DCEM2), considérant que l'étudiant avait suivi la séance si sa fiche remplie était rendue. L'assiduité était ainsi définie par le nombre de séances suivies par l'étudiant. Cette promotion comportant 100 étudiants, nous appellerons groupes A et B, respectivement, celui constitué par les 50 étudiants les plus assidus et celui constitué par les 50 étudiants les moins assidus.

Evaluation des performances des étudiants

A titre d'indicateurs de l'impact de l'intervention pédagogique étudiée (les SERAM), les moyennes des scores aux examens facultaires en DCEM2 et des rangs de classement à l'examen facultaire de DCEM4 et aux ENC ont été calculées et comparées en fonction de la variable assiduité.

Scores obtenus à une simulation d'ECN en DCEM2

Cet examen a été organisé, dans des conditions identiques à celles des ECN, en DCEM2, 8 mois après le début des SERAM, sur 5 dossiers rédigés selon les recommandations du CNCI, par les enseignants cliniciens ayant participé aux SERAM. Ces 5 dossiers, cotés sur 20 points, portaient sur les disciplines suivantes :

Dossier 1 : maladie infectieuse

Dossier 2 : cardiologie

Dossier 3 : sujet transversal sur tous les thèmes traités au cours du DCEM2

Dossier 4 : module 9 (maladies vasculaires)

Dossier 5 : hépato-gastro-entérologie.

L'étudiant n'ayant pas encore acquis des connaissances déclaratives suffisamment larges en DCEM2, la transversalité de ces dossiers n'était pas inter- mais intra-disciplinaire, c'est-à-dire que plusieurs thèmes d'une même discipline étaient abordés dans un même dossier.

Scores obtenus à l'examen de fin d'année en DCEM2

Cet examen portait sur six dossiers plutôt monodisciplinaires et un dossier transversal pluridisciplinaire.

Ces sept dossiers, cotés sur 20 points, portaient sur les disciplines suivantes :

Dossier 1 : médecine légale/santé publique

Dossier 2 : pathologie abdominale

Dossier 3 : pathologie thoracique

Dossier 4 : pathologie neuro-psychiatrique

Dossier 5 : maladies infectieuses

Dossier 6 : pathologie vasculaire

Dossier 7 : dossier multidisciplinaire

Rang de classement des étudiants à l'examen de fin d'année en DCEM4

Cet examen portait sur neuf dossiers, cotés sur 20 points, dont trois transversaux, deux de synthèse diagnostique, un portant sur le module 1 du programme des ECN, un de pharmacologie, un de médecine du travail et un de gériatrie. Il a été organisé en février 2004, trois mois avant les ECN.

Les dossiers de fin d'année ont été rédigés par un responsable de l'enseignement puis revus et corrigés par un comité de lecture vérifiant leur adéquation aux recommandations du CNCI.

Classements aux ECN

Les moyennes des rangs de classement (obtenus d'après le total des scores des neuf dossiers), calculées en fonction de la variable « assiduité », ont été comparées à l'issue de la session des ECN des 14 et 15 juin 2004.

Analyse statistique

Les comparaisons des moyennes des scores ont été effectuées par un test de l'écart réduit (paramétrique) pour les populations normales ou de Wilcoxon (non paramétrique) pour les distributions non normales. Les comparaisons de variables qualitatives ont été réalisées par le test du chi-deux. Les corrélations entre les rangs de classements aux différents examens ont été établies par le calcul du coefficient de corrélation de Pearson. Le seuil de significativité a été fixé à 5 %.

Des modèles linéaires mixtes (à pente et intercepte aléatoire) ont été utilisés pour étudier l'évolution des rangs entre le classement en première année du premier cycle des études médicales (PCEM1), en DCEM2, et DCEM4 et le classement aux ECN. Etant donné que les classements ont été effectués sur un nombre total différent d'étudiants, les quatre rangs ont été rapportés, pour cette analyse, sur une échelle de 100.

Résultats

Caractéristiques des groupes

Cent dix SERAM ont eu lieu pendant la période comprise entre le début du DCEM2 (quatrième année) et la fin du DCEM4 (sixième année), au cours de laquelle environ 350 dossiers couvrant l'ensemble du programme des ECN

ont été proposés aux étudiants. Trente SERAM ont eu lieu au cours du DCEM2. Les 50 étudiants les plus assidus ont suivi en moyenne 11,1 séances sur 30 possibles (minimum 6, maximum 30), tandis que les 50 étudiants les moins assidus ont suivi en moyenne 3,2 séances (minimum 0, maximum 5) (Tableau 1). Les étudiants du groupe A n'avaient pas été mieux classés à l'examen de PCEM1 que les étudiants du groupe B (rang moyen 52 vs 54 ; $p > 0,05$, test de l'écart réduit). Il n'existe pas de corrélation entre le rang de classement en PCEM1 et le nombre de SERAM suivies ($R = -0,13$; $p > 0,05$, test de corrélation de Pearson). Il n'y a pas de prépondérance d'étudiants primants c'est-à-dire reçus dès la première année de PCEM dans le groupe A ($p > 0,05$; test du chi-deux) (Tableau 1).

Performances des étudiants aux examens Scores obtenus à une simulation d'ECN en DCEM2 (Tableau 2)

Cet examen ne revêtait pas de caractère obligatoire. Certains étudiants n'ont pas répondu à tous les dossiers, ce qui explique les variations observées dans les effectifs. Les étudiants assidus étaient cependant environ deux fois plus nombreux. Les étudiants assidus ont obtenu une meilleure note au dossier 1 (maladies infectieuses) et au dossier 4 (module 9).

Scores obtenus aux examens de fin d'année en DCEM2 (Tableau 3)

Cet examen étant obligatoire, les effectifs étaient complets à

savoir 50 étudiants dans chaque groupe. Les étudiants assidus ont obtenu une meilleure note pour cinq des sept dossiers. Il existe une corrélation positive ($R = 0,63$; $p < 0,05$) entre les résultats à la simulation d'examen classant et les résultats en fin d'année en DCEM2.

Rangs de classements à l'examen de fin d'année en DCEM4 (Tableau 4)

Quatre-vingt onze étudiants se sont présentés à cette épreuve finale. L'effectif total des étudiants de cette promotion est cependant de 92. En effet, un étudiant a validé son année au Canada et n'a donc pas participé aux examens de fin d'année. Il est à noter que deux étudiants du groupe A et six du groupe B ont redoublé le DCEM2 ou le DCEM3 ce qui explique les différences d'effectifs observées entre le DCEM4 et le DCEM2. La moyenne des rangs de classement est meilleure de 10 places (41 vs 51) dans le groupe A, à la limite du seuil de significativité ($p > 0,07$: test de l'écart réduit).

Rangs de classement aux ECN (Tableau 4)

La moyenne des rangs de classement est significativement plus élevée dans le groupe B (1960 vs 1488, $p < 0,05$, test de l'écart réduit). Il existe une corrélation entre le nombre de SERAM suivies et le rang de classement aux ECN ($R = -0,32$; $p < 0,01$, coefficient de corrélation de Pearson) et entre le rang de classement en PCEM1 (première année du premier cycle des études médicales) et le rang aux ECN ($R = 0,37$; $p < 0,001$, coefficient de corrélation de Pearson).

Tableau 1 :
Caractéristiques des groupes d'étudiants étudiés

	Groupe A	Groupe B	p
Sexe ratio femmes/hommes	1,94	1,08	0,15*
Nombre de primants en PCEM1	14	6	0,07*
Rang moyen en PCEM1	55	52	0,6**
Nombre moyen de séances suivies en DCEM2	11,1	3,2	
Ecart-type	5	1	
Minimum - Maximum	6-30	0-5	

* test du chi-deux ; ** test de l'écart réduit

PCEM1 : 1^{re} année du 1^{er} cycle des études médicales

DCEM2 : 2^e année du 2^e cycle des études médicales

Groupe A : étudiants assidus aux séances d'entraînement au raisonnement médical

Groupe B : étudiants non assidus aux séances d'entraînement au raisonnement médical

Tableau 2 :
Moyenne des scores sur 20 à la simulation d'examen classant en DCEM2
(tous les dossiers ont la même échelle de cotation)

	Dossier 1	Dossier 2	Dossier 3	Dossier 4	Dossier 5
Groupe A					
<i>n</i>	45	45	44	41	39
m	12	14,3	8	9,9	10,8
Groupe B					
<i>n</i>	22	22	21	19	16
m	10	13,4	7,5	7,4	9,8
p (test de l'écart réduit)	0,01	NS	NS	0,01	NS

n = nombre d'étudiants ayant remis leur copie

m = moyenne des notes

NS = différence non significative

Groupe A : étudiants assidus aux séances d'entraînement au raisonnement médical

Groupe B : étudiants non assidus aux séances d'entraînement au raisonnement médical

Tableau 3 :
Moyennes des scores à l'examen de fin d'année en DCEM2

	D1	D2	D3	D4	D5	D6	D7
Groupe A (n = 50)	14,3	13,3	11,9	14,1	12,9	11,9	12,8
Groupe B (n = 50)	13,3	12,2	10,8	12,5	9,4	11	11,6
p	NS	0,05**	0,01*	0,005**	0,002*	0,04**	NS

DCEM2 : 2^e année du 2^e cycle des études médicales

D : Dossier ; * test de Wilcoxon ; ** test de l'écart réduit

NS = différence non significative

Tableau 4 :
Rangs de classement en DCEM4 et aux ECN

	Rang moyen en DCEM4	Rangs moyen aux ECN
Groupe A		
Effectif	48	48
Moyenne	41,6	1488
Ecart-type	27	1014
Groupe B		
Effectif	43	44*
Moyenne	51,4	1960
Ecart-type	24	874
p (Test écart réduit)	0,07	0,01

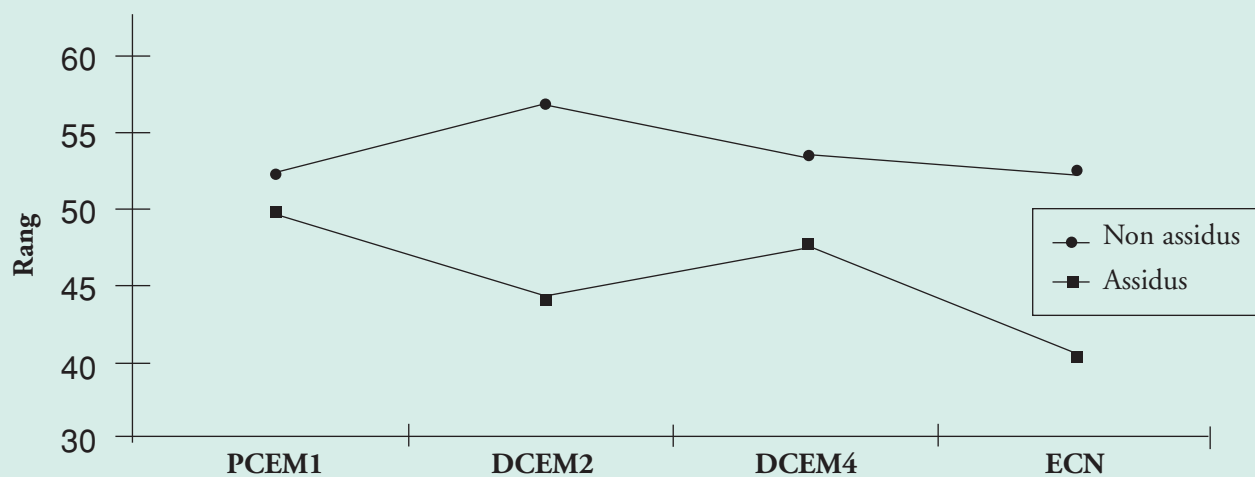
Groupe A : étudiants assidus aux séances d'entraînement au raisonnement médical
Groupe B : étudiants non assidus aux séances d'entraînement au raisonnement médical
ECN : épreuves classantes nationales
DCEM4 : 4^e année du 2^e cycle des études médicales
** un étudiant a validé sa dernière année du second cycle au Canada*

Tableau 5 :
Rangs moyens des étudiants en fonction de l'assiduité aux SERAM évaluée en DCEM2
(rapportés sur une échelle de 100)

	NON ASSIDUS		ASSIDUS	
	Moyenne	Ecart-type	Moyenne	Ecart-type
PCEM1	51,9	± 28,8	50,0	± 29,8
DCEM2	56,5	± 16,4	44,5	± 16,9
DCEM4	53,8	± 16,7	47,7	± 18,3
ECN	53,3	± 23,4	40,3	± 28,2

SERAM : séances d'entraînement au raisonnement médical
PCEM1 : 1^{re} année du 1^{er} cycle des études médicales
DCEM2 : 2^e année du 2^e cycle des études médicales
DCEM4 : 4^e année du 2^e cycle des études médicales
ECN : épreuves classantes nationales

Figure 1 :
Evolution des classements pour les étudiants assidus et non assidus



PCEM1 : 1^{ère} année du 1^{er} cycle des études médicales
DCEM2 : 2^e année du 2^e cycle des études médicales
DCEM4 : 4^e année du 2^e cycle des études médicales
ECN : épreuves classantes nationales

Tableau 6 :
Evolution des rangs entre le classement en PCEM1, DCEM2, DCEM4 et le classement aux ECN en fonction de l'assiduité

	IC 95 %	P
Intercept	(51,6 ; 65,7)	< 0,0001
Temps	(-3,5 ; -0,9)	0,24
Assiduité = non oui	(-15,9 ; -2,7)	0,0059

PCEM1 : 1^{ère} année du 1^{er} cycle des études médicales
DCEM2 : 2^e année du 2^e cycle des études médicales
DCEM4 : 4^e année du 2^e cycle des études médicales
ECN : épreuves classantes nationales
IC 95 % : intervalle de confiance à 95 %

Evolution des classements dans le temps

Les rangs moyens aux différents examens ont été rapportés sur une échelle de 100 (Tableau 5). L'évolution des classements dans le temps est représentée sur la figure 1. Ce type d'analyse statistique (Tableau 6) montre que les étudiants assidus ont des classements meilleurs que les étudiants non assidus quel que soit l'examen. Cependant, l'évolution des rangs dans le temps n'est pas significative.

Discussion

Ce travail est original par le fait qu'il se propose d'étudier l'impact à long terme d'une méthode d'intervention pédagogique formalisée, en suivant pendant trois ans une promotion d'étudiants ayant bénéficié des nouvelles mesures prises. Les SERAM constituent un format d'intervention pédagogique visant à favoriser chez l'étudiant le développement de certains aspects du raisonnement médical et, *in fine*, à préparer l'étudiant à un exercice professionnel de la médecine. Bien que les SERAM ne constituent pas en elles-mêmes un outil d'évaluation du raisonnement clinique tel que le test de concordance de script utilisé plus particulièrement dans les situations complexes¹⁶, la répétition de séances construites par des enseignants de disciplines variées, d'après des cas réels, constitue un réel outil d'évaluation formative. En effet, l'étudiant reçoit ainsi une rétroaction très régulière sur ses aptitudes au raisonnement médical et les enseignants peuvent, s'ils le souhaitent, suivre l'évolution de la promotion.

Même si le problème clinique constitue la base de ces séances et que celles-ci sollicitent un travail personnel important de la part des étudiants, il est clair qu'elles diffèrent de l'apprentissage par problème (APP), format d'intervention pédagogique habituellement utilisé dans le cadre du cursus préclinique mais visant à préparer l'apprentissage du raisonnement clinique^{20, 21, 22} ou encore des séances d'ARC¹⁴, généralement organisées dans le cadre des activités de l'externat. De plus, contrairement aux séances d'APP, les dossiers utilisés en SERAM sont construits de telle sorte que six à huit questions rédactionnelles courtes fassent suite à la présentation clinique. Ces questions sont posées de manière à tester prioritairement la justification d'une démarche clinique ou thérapeutique, en laissant la possibilité à l'étudiant de proposer plusieurs bonnes réponses.

Cette étude a cherché à évaluer l'efficacité des SERAM en tant que séances de préparation aux ECN. En tout état de cause, elle ne permet pas d'apprécier leur

impact à l'égard de la construction du raisonnement clinique chez nos étudiants, bien que cette préoccupation ait aussi été présente lors de la conception pédagogique des SERAM.

Avant d'en examiner les principaux résultats, il faut d'emblée préciser qu'il est très difficile de rapporter les effets observés à la seule mise en place de cet outil d'évaluation formative. Toute tentative de comparaison avec les années antérieures est malheureusement exposée à des biais statistiques majeurs tant les différences entre l'ancien concours de l'internat et les ECN sont importantes. Par ailleurs, il est parfaitement possible que les étudiants les plus assidus soient les plus laborieux et les plus motivés. Il aurait donc été très intéressant d'évaluer dans le même temps la motivation des étudiants pour comparer sur ce paramètre les deux groupes tout au long de leur cursus, tout en sachant que l'évaluation de la motivation intrinsèque est difficile tant elle est multifactorielle²³. A l'inverse, on peut également se demander si la mise en place des SERAM ne représente pas elle-même un facteur de motivation pour tous les étudiants, y compris ceux qui n'y participent pas.

La définition de l'assiduité telle qu'elle est utilisée ici n'est pas parfaite. Elle a été réalisée ponctuellement en DCEM2, en postulant que les deux groupes sont restés identiques pendant les deux années suivantes ce qui n'a probablement pas été le cas. Comme le soulignent Price et al.²⁴, en toute rigueur méthodologique, il aurait fallu randomiser les étudiants dans deux groupes, l'un suivant très régulièrement les SERAM et l'autre ne participant à aucune séance en maintenant cette obligation pendant trois ans. Un tel tirage au sort est bien sûr inconcevable dans le cadre d'une réforme institutionnelle pour des raisons pratiques et d'éthique institutionnelle évidentes. De plus, au-delà de son apparente légitimité « statistique », la pertinence d'essais contrôlés randomisés comporte de nombreuses limites pour ce qui concerne les travaux de recherche en éducation conduits en milieu naturel²⁵.

On peut cependant relever un certain nombre d'éléments valorisant l'assiduité aux SERAM : 1) les étudiants du groupe A n'étaient pas *a priori* « meilleurs » au sortir du PCEM1 ; 2) deux étudiants dans le groupe A ont redoublé contre six dans le groupe B et 3) on note une amélioration progressive et précoce (dès la fin du DCEM2) des performances des étudiants assidus aux différents examens puisqu'ils obtiennent un meilleur score sur deux dossiers à la simulation d'examen classant, cinq dossiers à l'examen de fin d'année en DCEM2 et une amélioration du rang de classement d'environ 500 places aux ECN ; groupes A et B confon-

du, il existe une corrélation entre le nombre de SERAM suivies et le classement aux ECN mais ce résultat doit être tempéré par la préexistence d'une corrélation entre le rang de classement en PCEM1 et celui aux ECN ; enfin, les termes de la définition de l'assiduité (les 50 % d'étudiants les plus assidus comparés aux autres) tendent probablement à minimiser l'effet éventuel des SERAM si l'on prend en considération la relative faible fréquentation de ces séances lors de la première année de mise en place (cf. *Tableau 1*).

Même si l'efficacité des programmes d'enseignement basés sur la résolution de problèmes n'est pas aujourd'hui complètement prouvée^{26,27}, il semble que la répétition d'une évaluation formative conduite avec des outils similaires à ceux de l'évaluation sommative²⁸ soit susceptible d'améliorer les performances de l'étudiant à deux niveaux bien distincts :

- l'acquisition d'une expertise par la confrontation à des problèmes complexes toujours différents, stimulant la génération précoce d'hypothèses, concept repris par Kassirer²⁹, permet des apprentissages contextualisés successifs et conduit à un rodage des compétences faisant appel à des connaissances conditionnelles. La gestion de la complexité des cas soumis permet également aux étudiants de se confronter à l'incertitude où plusieurs attitudes diagnostiques et/ou thérapeutiques restent possibles.
- Ce type de dispositif améliore en outre la capacité à traiter un format d'examen, l'étudiant se familiarisant avec la méthode de son évaluation finale.

Conclusion

Au cours du second cycle des études médicales, les facultés de médecine sont confrontées à un double enjeu difficile à concilier mais non incompatible : la formation de médecins compétents et la préparation des étudiants à des épreuves nationales classantes. L'initiation de séances d'entraînement au raisonnement médical tend à améliorer les performances des étudiants aux ECN, sous réserve des biais méthodologiques quasi-incontournables dans ce type d'étude. L'exigence d'avoir à préparer explicitement les étudiants à un examen national normatif aux modalités aussi univoques que ne le sont les ECN introduit en contrepartie des contraintes supplémentaires dans le cadre des activités d'enseignement, d'apprentissage et d'évaluation du raisonnement clinique. Cette expérience a permis d'amorcer, au sein de la faculté de médecine de Nantes, une réflexion sur les modalités docimologiques optimales permettant d'assurer un compromis acceptable entre performance aux ECN et qualité des apprentissages.

Contributions

Pierre Pottier a coordonné des séances d'entraînement au raisonnement médical (SERAM), recueilli les données et rédigé les versions successives du manuscrit. Bernard Planchon, vice-doyen chargé des études médicales, est co-instigateur des SERAM et responsable des enseignements du 2^e cycle. Jacques Barrier, responsable du département de développement pédagogique, a participé à l'interprétation des résultats et à la rédaction du manuscrit. Christelle Volteau, biostatisticienne, a participé à l'interprétation des résultats et a formulé des conseils sur l'analyse statistique. Alain Mouzard, doyen de la faculté de médecine de Nantes de 2000 à 2004, est instigateur des SERAM.

Annexe 1 : Recommandations du conseil scientifique du centre national des concours d'internat

Les dossiers devront permettre de :

1- Justifier une démarche diagnostique :

- analyser et hiérarchiser les données issues d'un examen clinique ;
- formuler des hypothèses diagnostiques initiales qui doivent venir à l'esprit à la lecture de l'observation (données cliniques et paracliniques exposées) et dire sur quels arguments elles sont évoquées ;
- déduire les examens complémentaires pertinents (biologie, imagerie, etc.) et les justifier.

2- Tester les connaissances physiopathologiques

3- Justifier le choix d'une attitude thérapeutique, parmi d'autres possibles, en insistant sur la gestion de l'incertitude, c'est-à-dire sur les bénéfices et risques éventuels qui seront explicités au malade et à sa famille.

4- Prendre en compte la dimension médico-légale

Rédaction de certificats, responsabilité médicale, déontologie, relations médecin – malade, déclaration obligatoire, rédaction d'une ordonnance (dans des situations fréquentes ou communes, rédiger avec précision une ordonnance incluant posologie, voie d'administration et durée du traitement ; pour les situations rares et/ou justifiant une prise en charge spécialisée, énoncer les principes thérapeutiques).

5- Prendre en compte la dimension sanitaire et sociale

Prévention, dépistage, attitude à adopter devant une épidémie, pathologie professionnelle, protection sociale, etc.

Annexe 2 : Exemples d'objectifs d'une séance

En DCEM2, chaque dossier était construit de manière à tester les connaissances de l'étudiant sur plusieurs thèmes de la même discipline.

75. Épidémiologie et prévention des maladies transmissibles : méthode de surveillance

- Préciser les bases de l'épidémiologie des maladies transmissibles et les mesures de surveillance et de prévention
- Déclarer une maladie transmissible.

80. Endocardite infectieuse

- Diagnostiquer une endocardite infectieuse.
- Argumenter l'attitude thérapeutique et planifier le suivi du patient.

81. Fièvre aiguë chez un malade immunodéprimé

- Identifier les situations d'urgence et planifier leur prise en charge.

85. Infection à VIH

- Informer et conseiller en matière de prévention de la transmission sanguine et sexuelle du VIH.
- Diagnostiquer une infection à VIH.
- Annoncer les résultats d'une sérologie VIH.
- Argumenter l'attitude thérapeutique et planifier le suivi du patient.

104. Septicémie

- Diagnostiquer une septicémie.
- Identifier les situations d'urgence et planifier leur prise en charge.
- Argumenter l'attitude thérapeutique et planifier le suivi du patient.

EN DCEM4

112. Réaction inflammatoire : aspects biologiques et cliniques ; conduite à tenir

- Expliquer les principaux mécanismes et les manifestations cliniques et biologiques de la réaction inflammatoire et les points d'impacts des thérapeutiques anti-inflammatoires.
- Argumenter les procédures diagnostiques devant un syndrome inflammatoire et/ou une VS élevée inexpliqués.

297. Anémie

Devant une anémie, argumenter les principales hypothèses diagnostiques et justifier les examens complémentaires pertinents :

les thèmes « orientations diagnostiques devant » sont pluridisciplinaires par définition.

323. Œdèmes des membres inférieurs

Devant l'apparition d'œdèmes des membres inférieurs, argumenter les principales hypothèses diagnostiques et justifier les examens complémentaires pertinents.

DCEM2 : 2^e année du 2^e cycle des études médicales

DCEM4 : 4^e année du 2^e cycle des études médicales

Annexe 3 :

Exemple de grille acceptant plusieurs réponses possibles

Une alternative acceptable :

Première solution

HBPM à dose curative en 1 injection (INNOHEP® 0,6 ml/j en SC)	5 points
Pendant 6 semaines	5 points
Car il s'agit d'une thrombose surale	3 points
Arrêt sous contrôle écho-doppler veineux	3 points
Avantages : évite le relais AVK, pas de surveillance INR	2 points
Inconvénients : injection quotidienne prolongée (risque de surdosage et inconfort)	2 points
Nécessité de surveiller les plaquettes 2 fois par semaine de J4 à J21	2 points
Activité anti-Xa au 10 ^e jour	2 points

0 à la question si plaquettes et notion de surdosage lié au traitement prolongé oubliés

Seconde solution

HBPM à dose curative en 1 injection	2,5 points
(INNOHEP® 0,6 ml/j en SC)	
Relais par PREVICAN® cp dès J1	2,5 points
Pendant 6 semaines	5 points
Car il s'agit d'une thrombose surale	3 points
Arrêt sous contrôle écho-doppler veineux	3 points
Surveillance INR régulière	1 point
Arrêt de l'INNOHEP quand l'INR est entre 2 et 3 à 48h d'intervalle	1 point
Inconvénients : risque hémorragique difficile à gérer, inconfort et stress lié à la nécessité de réaliser des INR	4 points
Avantage : confort lié à l'absence d'injection quotidienne	2 points

0 à la question si oublié des INR et oubli de la notion de risque hémorragique liés aux AVK

HBPM : Héparine de bas poids moléculaire ; AVK : anti-vitamine K ; INR : International normalized ratio

Références

1. Chamberland M. Comment exploiter les résultats de la recherche pour améliorer nos activités d'enseignement et d'apprentissage du raisonnement clinique ? *Pédagogie Médicale* 2005;6:197-9.
2. Eva KW. Ce que tout enseignant devrait savoir concernant le raisonnement clinique. *Pédagogie Médicale* 2005;6:225-34.
3. Nendaz M, Charlin B, Leblanc V, Bordage G. Le raisonnement clinique : données issues de la recherche et implications pour l'enseignement. *Pédagogie Médicale* 2005;6:235-54.
4. Elstein AS, Schwarz A. Clinical problem solving and diagnostic decision making : selective review of the cognitive literature. *BMJ* 2002;324:729-32.
5. Arocha JF, Wang D, Patel VL. Identifying reasoning strategies in medical decision making : a methodological guide. *J Biomed Inform* 2005;38:154-71.
6. Brooks LR, Norman GR, Allen SW. Role of specific similarity in a medical diagnostic task. *J Exp Psychol Gen* 1991;120:278-87.
7. Long WJ. Medical informatics : reasoning methods. *Artif Intell Med* 2001;71-87.

8. Kaufman DR, Patel VL. Problem solving in the clinical interview : a cognitive analysis of the performance of physicians, residents and students. *Teach Learn Med* 1991;3:6-14.
9. Barrows HS, Tamblyn R. Problem based learning : an approach to medical education. *Med Educ* 1974;84:2121-6.
10. Bordage G, Zacks R. The structure of medical knowledge in the memories of medical students and general practitioners : categories and prototypes. *Med Educ* 1984;18:406-16.
11. Lemieux M, Bordage G. Propositional versus structural semantic analyses of medical diagnostic thinking. *Cogn Science* 1992;16:185-204.
12. Kassirer J, Gorry G. Clinical problem-solving : a behavioural analysis. *Ann Intern Med* 1978;89:245-55.
13. Patel VL, Kaufman DR, Arocha JF. Emerging paradigms of cognition in medical decision-making. *J Biomed Inform* 2002;35:52-75.
14. Chamberland M. Les séances d'apprentissage du raisonnement clinique (ARC) : un exemple d'activité pédagogique contextualisée adaptée aux stages cliniques en médecine. *Ann Med Interne* 1998;149:479-84.
15. Charlin B, Bordage G, Van Der Vleuten C. L'évaluation du raisonnement clinique. *Pédagogie Médicale* 2003;4:42-52.
16. Charlin B, Brailovsky C, Leduc C, Blouin D. The diagnosis script questionnaire : a new tool to assess a specific dimension of clinical competence. *Adv Health Sci Educ Theory Pract* 1998;3:51-8.
17. Arrêté du 10 octobre 2000 modifiant l'arrêté du 4 mars 1997 relatif à la deuxième partie du deuxième cycle des études médicales. *Journal Officiel de la République Française*. [On-line]. Disponible sur : <http://www.legifrance.gouv.fr/>. NOR: MENS0002592A
18. Arrêté du 29 janvier 2004 relatif à l'organisation des épreuves classantes nationales anonymes donnant accès au troisième cycle spécialisé des études médicales. *Journal officiel de la République Française*. [On-line]. Disponible sur : <http://www.legifrance.gouv.fr/>. NOR: SANH0420335A
19. Epstein RJ. Learning from the problems of problem-based learning. *BMC Med Educ* 2004;4:1.
20. Colliver JA. Effectiveness of problem-based learning curricula. *Acad Med* 2000;75:259-66.
21. Smith PB, Verbeek AM, de Buissonjé CD. Problem based learning in continuing medical education : a review of controlled evaluation studies. *BMJ* 2002;324:153-6.
22. Chan DH, Leclair K. PBL and small group learning via internet among community family physicians. *MD comput* 1999;16:54-8.
23. Mish DA. Andragogy and medical education : are medical students internally motivated to learn ? *Adv Health Sci Educ Theory Pract* 2002;7:153-60.
24. Price EG, Beach MC, Gary TL, Robinson KA, Gozu A, Palacio A et al. A systematic review of the methodological rigor of studies evaluating cultural competence training of health professionals. *Acad Med* 2005;80:578-86.
25. Norman G. RCT = results confused and trivial: the perils of grand educational experiments. *Med Educ* 2003;37:582-4.
26. Moore GT, Black SD, Style CB. The influence of the new pathway curriculum on harvard medical students. *Acad Med* 1994;69:983-9.
27. Schmidt HG, Machiels-Bongaerts M, Herman H. The development of diagnostic competence : a comparison of a problem-based, an integrated and a conventional medical curriculum. *Acad Med* 1996;71:658-64.
28. Townsend AH, McIlvenny S, Miller CJ, Dunn EV. The use of an objective structured clinical examination (OSCE) for formative and summative assessment in a general practice clinical attachment and its relationship to final medical school examination performance. *Med Educ* 2001;35:841-6.
29. Kassirer JP. Teaching clinical medicine by iterative hypothesis testing : let's preach what we practice. *N Engl J Med* 1983;309:921-2.

Manuscrit reçu le 12 avril 2005 ; commentaires éditoriaux formulés aux auteurs le 3 octobre 2005, le 13 et le 19 septembre 2006 ; accepté pour publication le 20 septembre 2006.