



Mathieu Lorenzo

Département de médecine générale,
et centre de formation et de recherche
en pédagogie des sciences de la santé
(CFRPS), faculté de médecine de Strasbourg,
4, rue Kirschleger, 67085 Strasbourg

mlorenzo@unistra.fr

exercer 2021;176:375-9.

Liens et conflits d'intérêts : les auteurs déclarent
n'avoir aucun conflit d'intérêts en relation avec le
contenu de cet article.

Les liens d'intérêts éventuels de chacun
des auteurs sont disponibles sur le site :
www.transparence.sante.gouv.fr

Les problèmes à éléments clés

Un nouvel outil d'évaluation introduit par la réforme du deuxième cycle

*The key features examinations. A new assessment
tool introduced by the graduate medical school
reform*

INTRODUCTION

La réforme du deuxième cycle des études médicales (R2C) est entrée en vigueur pour les étudiants accédant à la première année du deuxième cycle à la rentrée 2020¹. Cette réforme tente d'inscrire davantage ce deuxième cycle dans une approche par compétences en visant notamment le développement du raisonnement clinique¹.

En cohérence avec ce dernier objectif, les modalités d'évaluation des apprentissages évoluent en intégrant deux outils relativement méconnus en France : les problèmes à éléments clés et le test de concordance de script (TCS).

Les problèmes à éléments clés ont été inventés dans les années 1980 à partir d'une volonté d'améliorer l'évaluation de la prise de décision des futurs médecins lors de l'examen qualifiant national du Canada^{2,3}. Ils semblent avoir atteint cet objectif avec une corrélation importante entre les scores aux problèmes à éléments clés et la qualité ultérieure de la pratique clinique⁴. De manière générale, les données de la littérature suggèrent que les problèmes à éléments clés sont un outil valide d'évaluation du raisonnement clinique^{4,5}. Ils sont largement utilisés dans les facultés de médecine en Amérique du Nord, aux Pays-Bas ou en Australie depuis plus de vingt-cinq ans⁵. Ils sont souvent désignés par leurs acronymes anglais KFE (pour *key features examinations*) ou KFP (pour *key features problems*).

Les problèmes à éléments clés explorent une sous-partie du processus de raisonnement clinique : la prise de décision clinique. Ils se présentent habituellement sous la forme de vignettes cliniques suivies d'une ou plusieurs questions. Les questions portent sur les éléments essentiels nécessaires pour diagnostiquer ou résoudre une situation clinique. Ce sont les « éléments clés » d'une situation clinique. Les problèmes à éléments clés, combinés dans le cadre de la R2C aux TCS, à des dossiers cliniques, des examens cliniques objectifs structurés (ECOS) et de la supervision en stage pourraient permettre une meilleure évaluation du processus de raisonnement clinique.

La discipline Médecine générale est officiellement en charge de certains des items du deuxième cycle. Le premier volume de l'ouvrage *Référence en médecine générale pour le 2^e cycle* rédigé sous la direction de Xavier Gocko propose 25 items abordés sous l'angle de la médecine générale⁶. Pour ces items en priorité, mais plus largement pour intégrer la vision de la médecine générale dans les évaluations du deuxième cycle, l'appropriation par les enseignants de la discipline des problèmes à éléments clés est nécessaire.

À la suite de l'article portant sur le TCS, l'objectif de cet article était de présenter le principe des problèmes à éléments clés et d'en détailler les principales modalités de conception.

PRINCIPES DES PROBLÈMES À ÉLÉMENTS CLÉS

La littérature en sciences de l'éducation montre que les experts n'utilisent que très peu d'éléments cliniques pour raisonner face à un patient^{2,7-9}. Ces quelques éléments clés utilisés par les experts sont contextuels : ils dépendent de la situation. En cohérence avec ces données, les problèmes à éléments clés se focalisent sur la capacité des étudiants à distinguer les éléments clés d'une situation clinique.

Dans un problème à éléments clés classique, une vignette textuelle détaillée est présentée à l'étudiant⁵. Celle-ci contient à la fois des éléments pertinents et non pertinents pour la résolution du cas⁵. Fréquemment, une première question va recueillir les hypothèses diagnostiques de l'étudiant face à la situation clinique⁵. Ce recueil s'effectue sous la forme d'une zone de texte libre en limitant fortement le nombre d'items possibles (ex : trois hypothèses au maximum). Ce recueil est peut-être hiérarchisé : quelle est l'hypothèse n° 1 ? la n° 2 ? etc.

D'autres types de questions peuvent demander à l'étudiant d'indiquer quelles sont les éléments clés de la vignette qui sous-tendent son hypothèse diagnostique principale⁵. L'étudiant choisit ainsi dans une longue liste des éléments présentés dans la vignette. Cette liste est souvent bien plus longue que dans une question à choix classique pour réduire les chances de réponses au hasard et les biais d'indication⁵.



Mme Haltenbach est une comptable retraitée de 68 ans. Elle se présente ce jour en consultation urgente à votre cabinet. Elle est très fatiguée depuis la veille et trouve qu'elle respire mal.

Vous la connaissez depuis trois ans. Elle a comme antécédents une fracture du poignet droit survenue il y a quatre ans, une appendicite opérée à l'âge de 28 ans, une hypertension artérielle depuis douze ans, une pneumopathie bactérienne du lobe supérieur droit il y a dix ans, une tendinopathie de la coiffe des rotateurs à droite à l'âge de 43 ans, une ablation d'un carcinome basocellulaire de l'oreille gauche il y a deux ans.

Son traitement habituel comprend de l'amlodipine 5 mg 1/j, du cholécalférol 100 000 UI (une ampoule tous les trois mois), du paracétamol 1 g en cas de douleurs.

À l'examen, sa température est à 38,5 °C, sa pression artérielle assise est à 162/88 mmHg au bras droit et à 159/84 mmHg au bras gauche. Sa saturation en oxygène est à 93 % en air ambiant, sa fréquence cardiaque à 96 batt/min.

L'auscultation cardiaque retrouve un souffle systolique aortique à 2/6, l'auscultation pulmonaire retrouve des crépitations des deux bases pulmonaires. Il n'y a pas de sibilants. Son abdomen est pléthorique, la palpation ne retrouve pas de défense ni de contracture. Il n'y a pas de masse abdominale.

Question 1 : Quelles sont vos trois hypothèses diagnostiques principales à ce stade ? Hiérarchisez.

1. _____
2. _____
3. _____

Question 2 : Quels sont les quatre éléments de l'énoncé les plus pertinents pour justifier votre hypothèse diagnostique principale parmi les propositions ci-dessous ?

Le sexe féminin	L'âge de 68 ans	La consultation en urgence	La fréquence cardiaque à 96 batt/min	La pression artérielle à 162/88 mmHg au bras droit
La pression artérielle à 159/84 mmHg au bras gauche	Les crépitations des deux bases pulmonaires	L'antécédent de fracture du poignet	L'antécédent d'appendicite	L'antécédent d'hypertension artérielle
L'antécédent de pneumopathie bactérienne	Le souffle systolique	L'antécédent de carcinome basocellulaire	Le traitement par amlodipine	Le traitement par cholécalférol
Le traitement par paracétamol	La température à 38,5 °C	L'absence de masse abdominale	La saturation à 93 %	Etc.

Question 3 : Quelles sont les éléments que vous voudriez particulièrement recueillir à l'interrogatoire en regard de votre hypothèse diagnostique principale ? Sélectionnez-en au maximum cinq.

Présence de frissons	Consommation quotidienne d'alcool	Expectorations purulentes	Hémoptysie	Palpitations
Suivi dermatologique de son basocellulaire	Présence de diarrhées	Perte d'odorat	Perte de goût	Insomnie
Présence d'acouphènes	Acuité visuelle habituelle	Douleurs dans les membres inférieurs	Heure précise de début des symptômes	Tabagisme
Douleur thoracique	Douleur pharyngée	Présence d'un rash cutané	Dysphagie associée	Etc.

Tableau 1 - Exemple de problème à éléments clés

Dans la même optique, les étudiants peuvent devoir indiquer dans une liste qui peut contenir des centaines de possibilités les quelques questions les plus pertinentes pour confirmer leur hypothèse diagnostique principale⁵. Il en est de même pour les manœuvres cliniques les plus pertinentes à proposer, les examens complémentaires pertinents, le traitement médicamenteux, etc. L'ensemble « vignettes + questions » est nommé « cas ». Un exemple de cas de problème à éléments clés selon ces principes est présenté dans le **tableau 1**.

Ce focus sur les éléments clés permettant de résoudre un problème est la différence principale entre les problèmes à éléments clés et les dossiers cliniques classiques. Ces derniers vont souvent évaluer des connaissances déclaratives, et non les éléments clés utilisés pour raisonner en situation de soins⁵. À titre d'exemple, une question du type « Quelles sont les molécules ayant l'indication pour le traitement de la thrombose veineuse profonde ? » explore un savoir théorique superficiel (demande de restituer des connaissances déclaratives).

À ce titre, des exemples de problèmes à éléments clés proposés par Montjotin-Gayton *et al.* dans le numéro 162 de cette revue sont questionnables¹⁰. Dans leur problème à éléments clés de l'item 47 « Puberté normale et pathologique », la première question est « À partir de quel âge parle-t-on de retard pubertaire chez le garçon ? ». Cette question est un exemple typique d'évaluation de connaissance déclarative. Il n'y a besoin d'aucun contexte clinique pour y répondre car la réponse n'en dépend pas. À l'inverse, pour l'item 44 « Suivi d'un nourrisson et d'un enfant normal », la question « Quel diagnostic le plus probable concernant l'œil gauche de Y suspectez-vous ? » suivant une vignette traitant d'une anomalie de l'œil chez un nourrisson de 12 mois semble un bon exemple de question de problème à éléments clés.

Une limite importante des problèmes à éléments clés est qu'ils ne

peuvent pas évaluer seuls l'intégralité du processus de raisonnement clinique^{11,12}. Les problèmes à éléments clés n'évaluent notamment pas correctement la justification diagnostique ou la représentation du problème¹². Seule une combinaison de plusieurs outils et méthodes d'évaluation du raisonnement clinique permet de correctement l'évaluer¹².

MODALITÉS DE CONCEPTION DES PROBLÈMES À ÉLÉMENTS CLÉS

La rédaction des problèmes à éléments à clés nécessite plusieurs heures et plusieurs rédacteurs¹². En effet, il est conseillé de réunir au minimum trois experts du domaine pour déterminer par consensus quels sont les éléments clés de la situation¹².

La première étape est de déterminer les objectifs de l'évaluation. S'agit-il d'une évaluation formative (visant à guider les apprentissages ultérieurs) ou d'une évaluation sommative (visant à vérifier que les apprentissages attendus ont eu lieu) ? Dans ce dernier cas, partir des objectifs d'apprentissage du référentiel de formation puis utiliser une table de spécification semble utile pour s'assurer d'une bonne validité de contenu des problèmes à éléments clés¹¹. Une table de spécification est un tableau croisant les objectifs d'apprentissage avec le contenu d'une évaluation¹³. Elle sert à s'assurer que l'évaluation couvre bien tous les éléments importants du programme (c'est la validité de contenu). Un exemple de table de spécification est donné dans le **tableau 2**. Le lien avec le rang des connaissances de la R2C (A, B et C)

n'est pas établi à ce jour. Dans la R2C, les connaissances de rang A sont à acquérir par tous les étudiants de 2^e cycle. Les connaissances de rang B sont des connaissances d'une spécialité exigibles au premier jour du DES. Les connaissances de rang C sont des éléments de spécialité uniquement enseignées en troisième cycle.

Les problèmes à éléments clés sont une des modalités de l'évaluation des connaissances tout en visant à documenter le développement d'une compétence (le raisonnement clinique). Il faudra choisir une des situations de départ (SDD) du nouveau référentiel et faire le lien avec les rangs de connaissances à l'aide des fiches LiSA (livret de suivi des apprentissages). Les fiches SDD et LiSA sont en cours de finalisation au niveau national à ce jour. Dans le cas d'une évaluation formative, ces considérations de validité de contenu sont moins importantes, et l'accent est mis sur la rétro-information offerte à l'étudiant à la suite de la réalisation des problèmes à éléments clés¹³. C'est cette rétro-information qui permet à l'étudiant de cerner les apprentissages à réaliser pour la suite. Il semble donc utile de prendre un temps dédié visant à expliciter le processus de raisonnement clinique de l'expert sous-tendant la résolution des problèmes à éléments clés pour modéliser celui de l'étudiant¹⁴.

Pour écrire les cas, le plus pertinent est de partir d'une situation réelle vécue par l'un des membres du comité de rédaction. Dans cette étape, il est utile de contextualiser la vignette¹⁵. Le fait de donner un nom (fictif) au patient, d'ajouter une profession, des détails personnels, tend à majorer l'engagement des étudiants dans une

utilisation formative des problèmes à éléments clés¹⁵. Il est dans tous les cas indispensable de présenter les valeurs brutes (sans interprétation) dans la vignette et dans les questions^{4,5,15}. Il est conseillé d'éviter d'écrire « il a de la fièvre » mais plutôt « sa température est à 38,5 °C ». Comme dans la vraie vie, cette étape d'interprétation des données permet de mieux discriminer les étudiants en difficulté de raisonnement clinique^{14,16}. La vignette présentée en exemple dans le **tableau 1** reprend ces principes.

Selon les objectifs de l'évaluation, le comité de rédaction propose de s'intéresser à la génération d'hypothèses diagnostiques, à l'investigation clinique, à la thérapeutique.

Vient ensuite le choix de la modalité de réponse : question ouverte à réponse courte ? Question à choix multiple ? Choisissez le nombre de réponses attendues pour chaque question et le nombre maximal de propositions possibles (ce nombre peut être légèrement supérieur au nombre de réponses attendues). Dans l'exemple présenté dans le **tableau 1**, il est demandé à l'étudiant d'indiquer trois hypothèses diagnostiques. Trois hypothèses étaient attendues et valaient des points. Une autre possibilité aurait été d'indiquer à l'étudiant de renseigner « trois hypothèses au maximum » et de n'attribuer des points qu'à deux hypothèses. Pour les questions à choix multiple, déterminez le nombre de modalités que vous allez proposer au choix. Il est habituel que ce nombre soit au moins égal au double, voire au triple, du nombre maximal de réponses possibles pour limiter les réponses au hasard^{5,12}.

	Reconnaître les symptômes et signes cliniques [...]	Prescrire les examens adaptés [...]	Prendre en charge les motifs de consultation les plus fréquents [...]	Donner les informations adaptées [...]	Etc.
Nombre de cas	3	2	4	1	...
Nombre de questions	4	4	5	2	...

Tableau 2 - Exemple d'une table de spécification de problèmes à éléments clés pour l'item 22 « Grossesse normale »



La phase cruciale est celle de la détermination par consensus des éléments clés. Quelles sont les hypothèses diagnostiques les plus probables ? Quels sont les éléments de l'énoncé les plus contributifs ? etc. C'est cette phase qui prend le plus de temps.

Dans les questions à choix multiple, il faut choisir les distracteurs (les options de réponses proposées en sus des éléments clés attendus). Les meilleurs distracteurs sont les plus plausibles. Une méthode efficace semble être d'utiliser les erreurs les plus fréquentes des étudiants. Pour cela, une question sous la forme de question ouverte courte peut être posée préalablement à un groupe d'étudiants. Utilisez ensuite les réponses erronées pour créer vos distracteurs et transformer cette question en question à choix pour le groupe d'étudiants suivant.

Pour les questions ouvertes courtes, il est conseillé de déterminer par consensus une grille de correction comprenant les éléments clés attendus, leurs synonymes et équivalences acceptés. La fidélité semble meilleure en recourant à une correction automatisée par ordinateur pour ces questions⁴.

Le nombre de cas à créer dépend des objectifs de l'évaluation. Dans une évaluation sommative ou certificative à fort enjeu, il faut une quarantaine de cas pour atteindre une fidélité suffisante^{5,12}.

Un tel examen de 40 cas aura une durée d'environ quatre heures². Il semble plus pertinent dans un même temps de limiter à deux ou trois le nombre de questions par cas pour poser plus de cas¹⁷. La performance en termes de raisonnement clinique étant contextuelle (être expert dans une situation ne prédit pas d'être expert dans une autre), son évaluation est davantage valide avec une large variété de cas traitant de multiples situations cliniques différentes^{17,18}. Dans le cas d'un examen à fort enjeu comme le sera l'examen national de fin de deuxième cycle, ces recommandations de nombre de cas et de nombre de questions par cas devraient logiquement s'appliquer.

Le mode de notation présente également des spécificités. Les recherches ont montré qu'il était préférable d'attribuer le même nombre de points à chaque cas au sein d'un ensemble, à chaque question au sein d'un cas et à chaque élément de réponse au sein d'une question^{4,5}. Le calcul du score est une succession de calculs de moyennes (**encadré**).

1. L'étudiant a indiqué 2/3 des hypothèses diagnostiques attendus à la question n° 1. Il a correctement renseigné 4/4 éléments à la question n° 2 et 3/5 à la question n° 3.
Son score à ce cas est
 $(2/3 + 4/4 + 3/5)/3 = 75 \%$

2. Sur un ensemble de quatre cas, ses scores sont 75 %, 86 %, 52 % et 45 %.
Son score global est
 $(0,75 + 0,86 + 0,52 + 0,45) / 4 = 64,5 \%$.

Encadré - Exemple de calcul de scores dans un problème à éléments clés

Des modes de calcul plus complexes sont décrits comme des attributions de points négatifs, ou des pénalités en cas de nombre trop important de réponses sont également décrites sans que leurs valeurs ajoutées soient évidentes^{4,5}.

CONCLUSION

L'introduction des problèmes à éléments clés dans la R2C est une opportunité de mieux évaluer le raisonnement clinique des étudiants en médecine. Cet article pourrait favoriser l'appropriation des problèmes à éléments clés par les enseignants universitaires de médecine générale. Ainsi, une nouvelle dynamique pourrait s'enclencher afin d'alimenter les bases de questions nationales avec des problèmes à éléments clés en lien avec les soins premiers. ◆

Résumé

La réforme du deuxième cycle des études médicales en France tente d'inscrire davantage celui-ci dans une approche par compétences en visant notamment le développement du raisonnement clinique. En cohérence avec ce dernier objectif, un nouvel outil d'évaluation a été introduit : les problèmes à éléments clés. L'objectif de cet article est de présenter le principe des problèmes à éléments clés et d'en détailler les principales modalités de conception.

Les problèmes à éléments clés explorent une sous-partie du processus de raisonnement clinique : la prise de décision clinique. Ils se présentent habituellement sous la forme de vignettes cliniques suivies d'une ou plusieurs questions. Les questions portent sur les éléments essentiels nécessaires pour diagnostiquer ou résoudre une situation clinique. Ce sont les « éléments clés » d'une situation clinique. Ce travail propose une méthode pour créer des problèmes à éléments clés à partir d'une situation vécue. Il est conseillé de se réunir à trois experts du domaine pour déterminer par consensus les éléments clés de la situation. L'exploration de la phase précoce du raisonnement clinique au travers de la génération initiale d'hypothèses diagnostiques est pertinente dans un problème à éléments clés.

→ **Mots-clés MESH** : éducation ; enseignement spécialisé en médecine ; enseignement ; pédagogie.

Summary

The reform of the graduate medical school in France tries to align more with a competency-based approach by focusing in particular on the development of clinical reasoning. In line with this last objective, a new assessment tool is introduced: the key features problems. The objective of this article is to present the principle of key features problems and to detail how to design it.

Key features problems explore a sub-part of the clinical reasoning process: decision making. They usually take the form of clinical vignettes followed by one or more questions. The questions focus on the essential elements needed to diagnose or resolve a clinical situation. These are the « key features » of a clinical situation. Determining by consensus of three or more people the key features of a situation from the trainers' clinical experience is recommended. Exploring the early phase of clinical reasoning and especially the generation of diagnostic hypotheses is often a good option.

→ **Keywords MESH**: education; education; medical; graduate; teaching; pedagogy.

Références

1. République française. Arrêté du 2 septembre 2020 portant modification de diverses dispositions relatives au régime des études en vue du premier et du deuxième cycle des études médicales et à l'organisation des épreuves classantes nationales.
2. Page G, Bordage G. The medical council of Canada's key features project: a more valid written examination of clinical decision-making skills. *Acad Med J* 1995;70:104-10.
3. Hatala R, Norman GR. Adapting the key features examination for a clinical clerkship. *Med Educ* 2002;36:160-5.
4. Bordage G, Page G. The key-features approach to assess clinical decisions: validity evidence to date. *Adv Health Sci Educ* 2018;1-32.
5. Hrynychak P, Takahashi SG, Nayer M. Key-feature questions for assessment of clinical reasoning: a literature review. *Med Educ* 2014;48:870-83.
6. Gocko X. Références en médecine générale pour le deuxième cycle - 25 Items. Paris : CNGE, 2019. Disponible sur : <https://www.exercer.fr/librairie/produits/produit/70> [consulté le 1 septembre 2021].
7. Pelaccia T, Tardif J, Tribby E, Charlin B. An analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. *Med Educ Online* 2011;16.
8. Pelaccia T, Tardif J, Tribby E, et al. How and when do expert emergency physicians generate and evaluate diagnostic hypotheses? A qualitative study using head-mounted video cued-recall interviews. *Ann Emerg Med* 2014.
9. Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ* 2005;39:98-106.
10. Montjotin-Gayton L, Rodrigues Sousa Maguin M, Plotton C, Gocko X. Nouvelle docimologie pour le deuxième cycle des études médicales. *exercer* 2020;185-6.
11. Bertrand C, Dory V, Pelaccia T, et al. Choisir un outil d'évaluation. In: Comment (mieux) former et évaluer les étudiants en médecine et en sciences de la santé ? Bruxelles : De Boeck Supérieur, 2016:33-44.
12. Daniel M, Rencic J, Durning SJ, et al. Clinical reasoning assessment methods: a scoping review and practical guidance. *Acad Med* 2019;94:902-12.
13. Bertrand C, Dory V, Pelaccia T, Charlin B, Hodges B. Les principes généraux de l'évaluation. In: Comment (mieux) former et évaluer les étudiants en médecine et en sciences de la santé ? Bruxelles : De Boeck Supérieur, 2016:343-55.
14. Faucher C, Pelaccia T, Nendaz M, Audétat MC, Charlin B. Un professionnel de santé qui résout efficacement les problèmes : le raisonnement clinique. In: Comment (mieux) former et évaluer les étudiants en médecine et en sciences de la santé ? Bruxelles : De Boeck Supérieur, 2016:33-44.
15. Cohen DA, Newman LR, Fishman LN. Twelve tips on writing a discussion case that facilitates teaching and engages learners. *Med Teach* 2017;39:147-52.
16. Audétat MC, Laurin S. Supervision of clinical reasoning: methods and a tool to support and promote clinical reasoning. *Can Fam Physician* 2010;56:e127-129, 294-6.
17. Norman G, Bordage G, Page G, Keane D. How specific is case specificity? *Med Educ* 2006;40:618-23.
18. Croskerry P. Context is everything or how could I have been that stupid? *Healthc Q Tor Ont* 2009;12:e171-176.