




Disponible en ligne sur
 ScienceDirect
 www.sciencedirect.com

Elsevier Masson France

 www.em-consulte.com



Article original

Mise en œuvre, en situation professionnelle, des recommandations de la HAS, par la méthode du test de concordance de script combinée à l'analyse actualisée de la littérature. Application à la polyarthrite rhumatoïde

Xavier Le Loët^{a,*}, Thierry Lequerré^a, Alain Cantagrel^b, Louis Sibert^c

^a Inserm U 905, IFR 23, service de rhumatologie, CHU, hôpitaux de Rouen, 1, rue de Germont, 76031 Rouen cedex, France

^b Service de rhumatologie, hôpital Larrey, CHU de Toulouse, TSA 30030, 31059 Toulouse, France

^c EA 4308, IFR 23, service d'urologie, CHU, hôpitaux de Rouen, 1, rue de Germont, 76031 Rouen cedex, France

INFO ARTICLE

Historique de l'article :

Accepté le 8 juillet 2011

Disponible sur Internet le 23 septembre

2011

Mots clés :

Test de concordance de script

Revue de la littérature

Recommandations Haute Autorité de santé

Polyarthrite rhumatoïde

Évaluation des pratiques professionnelles

RÉSUMÉ

Objectifs. – Décrire et évaluer la faisabilité d'un processus original d'évaluation des pratiques professionnelles, mis en œuvre lors d'une formation professionnelle annuelle dénommée « Rencontres d'experts en rhumatologie ».

Méthodes. – Celui-ci, fondé sur une approche pédagogique nouvelle, le test de concordance de script, associée à la revue actualisée de la littérature, permet d'évaluer les pratiques, in vitro, en analysant les « zones d'incertitude » rencontrées dans des recommandations, ici, celles de la Haute Autorité de santé consacrées à la polyarthrite rhumatoïde. Après sélection de ces « zones d'incertitude » par un Conseil scientifique universitaire, environ 80 rhumatologues libéraux et hospitaliers sont interrogés sur des vignettes cliniques élaborées, conformément à la technique du test de concordance de script, et portant sur des situations équivoques. Une revue actualisée de la littérature concernant les sujets abordés est alors présentée et discutée. Puis, les participants à cette formation sont à nouveau interrogés sur ces mêmes vignettes. L'impact de cette revue de la littérature/discussion est enfin mesuré.

Résultats. – Cette approche originale, jamais utilisée pour évaluer les pratiques professionnelles, s'est avérée faisable et bien acceptée par des rhumatologues libéraux et hospitaliers, comme l'atteste le bon niveau de satisfaction.

Discussion. – Cette première expérience positive incite à utiliser cet outil à une plus grande échelle pour l'évaluation des pratiques professionnelles.

© 2011 Société Française de Rhumatologie. Publié par Elsevier Masson SAS. Tous droits réservés.

1. Introduction

Depuis quelques années ont été mises en place aux plans international et national des recommandations concernant la prise en charge des pathologies de l'appareil locomoteur [1–5]. En France, nous disposons de plusieurs recommandations professionnelles élaborées par la Haute Autorité de santé (HAS) [5]. Ces recommandations servent de référence à l'évaluation des pratiques. Cette évaluation consiste à mesurer, a posteriori, d'éventuels écarts entre la pratique et ces recommandations. Elle rencontre toutefois certains obstacles : les méthodes préconisées (audit, chemin clinique, revue de morbidité-mortalité, études de dossiers, réseaux de santé...) sont assez complexes, parfois difficiles à généraliser.

Les recommandations professionnelles s'appuient essentiellement sur les données scientifiques recueillies dans des groupes

homogènes de patients. Il est assez rare que les problèmes rencontrés au quotidien puissent être résolus par la simple application de ces recommandations. En pratique (in vivo), la prise de décision intègre des données autres que les connaissances fondées sur les preuves : caractéristiques du patient, comorbidités, environnement psychosocial, etc. . .

Ainsi, face au patient, une part significative de la compétence médicale relève de l'aptitude à composer avec l'incertitude.

L'évaluation de la mise en application des référentiels professionnels devrait utiliser des méthodes prenant en compte cette incertitude. Nous rapportons une approche originale d'évaluation in vitro des pratiques professionnelles mise au point pour les journées annuelles de formation dénommées « Rencontres d'experts en rhumatologie » (RER). Elle a été mise en œuvre à propos des recommandations de la HAS portant sur la polyarthrite rhumatoïde (PR) [5]. Pour ce faire, nous avons utilisé deux outils : le test de concordance de script (TCS), outil validé depuis plus de dix ans, qui vise à évaluer le raisonnement en interprétant des données cliniques en contexte d'incertitude [6]. Ce test est composé de cas cliniques

* Auteur correspondant.

Adresse e-mail : xavier.le-loet@chu-rouen.fr (X. Le Loët).

comportant une ambiguïté. Les participants sont invités à réagir, individuellement, face à ces situations [6].

Par ailleurs, la littérature scientifique qui demeure la base de la connaissance médicale [7]. Elle a servi à élaborer les recommandations de la HAS, notamment celles consacrées à la PR. Depuis leur diffusion, en 2007, la littérature concernant ce thème s'étant enrichie, une revue actualisée de celle-ci, en lien avec les cas cliniques précités, a donc été réalisée et présentée aux participants. Afin de mesurer l'impact de cette revue de la littérature sur leurs décisions, les réponses aux TCS ont été mesurées avant et après l'exposé de la littérature.

L'objectif de cet article est de décrire le processus général mis en œuvre, depuis deux ans, au cours des RER. Nous avons voulu vérifier :

- la faisabilité d'une telle approche pour mesurer l'écart entre les recommandations de la HAS et la décision prise in vitro face à des cas cliniques réalistes, selon que les participants ont ou non connaissance de la littérature actualisée ;
- l'acceptabilité de la part des rhumatologues participant à cette démarche.

Les résultats concernant la mesure de l'impact de la connaissance de la littérature sur les décisions prises face à des situations cliniques définies dans des TCS fera l'objet d'une autre publication.

Il s'agit, à notre connaissance, de la première expérience publiée ayant trait à l'utilisation du TCS pour évaluer l'impact réel des recommandations professionnelles.

2. Méthodes

2.1. Rencontres d'experts en rhumatologie : le Comité scientifique ; organisation générale

Depuis 2004 ont été mises en place, à l'échelon national, ces RER. Cet événement pédagogique annuel rassemble, en un même lieu, pendant deux jours, des rhumatologues libéraux et hospitaliers sur un thème ayant trait à la prise en charge des principaux rhumatismes inflammatoires. Un Conseil scientifique indépendant de dix rhumatologues universitaires français, experts en la matière, est chargé de préparer ces RER pendant les neuf mois précédents.

De 2004 à 2008, ces rencontres ont eu pour objet l'élaboration de recommandations pratiques concernant la prise en charge de la PR et des spondylarthropathies [8]. Entre temps, la publication de recommandations institutionnelles internationales et françaises, en particulier de la HAS, rendait vaine la poursuite de l'élaboration de nouvelles recommandations qui ne pouvaient disposer d'un label « officiel ». Aussi, le Conseil scientifique a-t-il souhaité faire évoluer les objectifs de ces rencontres. Ainsi, en 2009 et 2010, ceux-ci sont devenus :

- l'analyse des décisions « spontanées » (avant présentation de la revue actualisée de la littérature) de rhumatologues, dans des situations dites « zones d'incertitude », présentes dans les recommandations de la HAS sur la PR ;
- puis, la mesure de l'impact de la connaissance de la littérature actualisée et d'une discussion entre ces professionnels, sur les décisions initiales « spontanées ». Au cours de ces deux étapes, l'outil utilisé était le TCS.

2.2. Recommandations de la Haute Autorité de santé

Les recommandations professionnelles de la HAS concernant la PR, publiées en septembre 2007 [5], rassemblent les points essentiels du diagnostic et de la prise en charge. Dans leur version

Propositions de stratégie thérapeutique de deuxième intention : pour un patient sous méthotrexate, insuffisant à dose maximale tolérée pendant au moins 3 mois.

La polyarthrite rhumatoïde est active, ou évolutive,

- soit avec un score de DAS >5,1,
- soit avec un score de DAS 28 >3,2 et une corticodépendance,
- soit avec la présence ou la **progression des lésions structurales à l'imagerie**.

Prescrire en fonction du niveau d'activité de la polyarthrite rhumatoïde,

âge, comorbidité, souhait du patient



Méthotrexate associé à un anti TNF alpha

ou adalimumab ou etanercept ou infliximab.

Zone d'incertitude: quelle est la définition de progression des lésions structurales à l'imagerie ?

Fig. 1. Exemple de « zone d'incertitude » dans un algorithme des recommandations professionnelles de la Haute Autorité de santé concernant la polyarthrite rhumatoïde [4].

abrégée – celle habituellement utilisée par les rhumatologues – ces recommandations sont présentées sous forme d'algorithmes simples, de lecture aisée. Elles envisagent la majorité des situations « théoriques » susceptibles d'être rencontrées en pratique quotidienne. Une version longue et complète de ces recommandations avec l'argumentaire y afférent est aussi disponible, par l'Internet (www.has-sante.fr).

2.3. Identification des « zones d'incertitude » des recommandations de la Haute Autorité de santé

Les membres du Comité scientifique ont sélectionné, par une méthode Delphi [9], les situations rencontrées dans les algorithmes de la HAS dont l'interprétation ne leur paraissait pas univoque. Ils pouvaient aussi retenir celles où les connaissances scientifiques avaient évolué depuis 2006, terme de l'analyse de la littérature par les experts de la HAS. Ainsi, ont été sélectionnés des points équivoques portant sur le diagnostic clinique, les examens complémentaires et la thérapeutique. Un exemple de situation équivoque ou « zone d'incertitude » figure dans le Fig. 1 ; il s'agit ici d'une recommandation thérapeutique en cas de « progression des lésions structurales à l'imagerie ». L'absence de définition précise de la « progression » rend cette recommandation équivoque. Une « zone d'incertitude » était sélectionnée chaque fois qu'elle recueillait au moins 70% des suffrages. Ces situations ont fait l'objet, par consensus, d'un bref libellé posant le problème de façon pratique.

2.4. Tests de concordance de script appliqués aux « zones d'incertitude » : rationnel et description du concept

Le principe et la méthodologie du TCS ont été largement décrits dans la littérature [10–13]. Schématiquement, il s'agit d'un outil d'évaluation du raisonnement clinique fondé sur la théorie des scripts et le développement de l'expertise médicale, issue de la psychologie cognitive [12–15]. Cette théorie postule que, face à une situation donnée, le médecin mobilise des réseaux de connaissance préétablis dénommés « scripts » qui s'enrichissent au fur et à mesure de son expérience clinique. L'organisation de ces réseaux de connaissances dépend de l'expérience personnelle du praticien. Cela explique que, quoique adoptant des modes de raisonnement parfois très différents, plusieurs cliniciens expérimentés, soumis aux mêmes situations complexes ou mal définies, adoptent finalement la même décision.

Classiquement, le TCS vise à mesurer le processus de prise de décision en contexte d'incertitude, en comparant les réponses des participants à celles d'un panel de cliniciens expérimentés dans le domaine testé. Le format du test est fait de cas cliniques brièvement décrits sous forme d'une « vignette ». Chaque situation contient suffisamment d'informations pour être crédible, mais pas assez pour pouvoir être résolue telle quelle. D'autres informations sont nécessaires.

Chaque situation, présentée de façon équivoque est suivie d'une série de questions d'ordre diagnostique, d'investigation ou de traitement. Chaque question est constituée de trois parties : la première est une hypothèse diagnostique, d'investigation ou de traitement pertinente, la deuxième est une nouvelle information absente du scénario initial (information clinique ou paraclinique). L'exercice demandé (troisième partie) à chaque participant est de mesurer l'impact que produit sur l'hypothèse initiale, cette nouvelle information. Comme il s'agit de porter un jugement d'ordre qualitatif et non quantitatif, cette dernière partie est présentée sous forme d'une échelle de Likert en 5 points. Un exemple de TCS est présenté dans le [tableau 1](#).

Dans sa version originale, dédiée à des médecins en formation, le système de notation du test est fondé sur la méthode des scores compilés [15]. Parallèlement aux étudiants, un groupe de médecins spécialistes dénommé « panel » effectue aussi le test, individuellement. Les réponses sont prises en compte pour définir la pondération affectée à chaque question. En d'autres termes, il s'agit de mesurer la « concordance » des jugements portés par les participants avec ceux de cliniciens expérimentés. L'objectif est bien de comparer les stratégies de raisonnement pour lesquelles il n'y a pas qu'une seule bonne réponse. Dans notre étude, les participants étant des professionnels avertis en matière de prise en charge de la PR, nous n'avons pas évalué la concordance de leurs jugements avec ceux d'experts, par exemple avec les membres du Conseil scientifique. En revanche, nous avons comparé leur décision, selon qu'ils avaient ou non connaissance de la revue actualisée de la littérature consacrée à ce sujet.

2.5. Modalités de la revue de la littérature

La revue de la littérature a été réalisée conformément aux recommandations actuelles. Schématiquement, pour chacun des sujets sélectionnés, ont été interrogées les bases de données Medline, Embase et Cochrane à la recherche d'articles publiés dans le domaine concerné jusqu'au mois d'avril, la réunion d'experts se déroulant fin mai-début juin. Les mots-clés étaient sélectionnés, soit à partir d'un thesaurus (MeSH pour Medline et Emtree pour Embase), ou à partir des titres et des résumés. Si nécessaire, les opérateurs AND, OR, AND NOT ont été utilisés pour combiner les mots-clés. Nous avons aussi sélectionné les travaux concernant les sujets étudiés, publiés sous forme de résumés au cours des congrès annuels de l'American College of Rheumatology, de l'European League Against Rheumatism et de la Société française de rhumatologie des deux dernières années. L'analyse était limitée aux articles publiés en anglais ou en français dans des revues à comité de lecture.

Cette analyse systématique de la littérature a été confiée à trois rhumatologues juniors (bibliographes) préalablement formés à cette méthodologie.

2.6. Description de la procédure et du calendrier

2.6.1. Phase préparatoire (septembre à mai de l'année suivante)

Celle-ci ne concerne que les membres du Conseil scientifique et les trois bibliographes :

- étape 1 : sélection des « zones d'incertitude » des algorithmes de la HAS jugées comme pertinentes par les experts du Comité scientifique ;
- étape 2 : regroupement des « zones d'incertitude » sélectionnées en trois domaines cohérents, afin que l'exercice pédagogique soit réalisé dans trois ateliers. Ceux-ci réunissent, en parallèle, environ 25 participants. Pour chaque domaine, est désigné un universitaire référent assisté d'un junior bibliographe ;
- étape 3 : en plénière du Conseil scientifique, définition précise du périmètre de chaque domaine et formulation des libellés correspondant aux « zones d'ombre » sélectionnées ;
- étapes 4 et 5 (contemporaines) : en plénière, découlant de l'étape précédente, le Conseil scientifique délimite le champ de la littérature à analyser par les bibliographes et élabore les vignettes cliniques présentées sous forme de TCS pour explorer les domaines prédéfinis. Pour la construction des TCS, nous avons utilisé les recommandations des concepteurs de la méthode [13,15]. Un exemple de vignette est présenté dans le [tableau 2](#) ;
- étape 6 : validation, par l'ensemble du Conseil scientifique, de l'analyse actualisée de la littérature et des vignettes au cours de deux réunions.

2.7. Rencontres d'experts en rhumatologie (fin mai – début juin)

Ces RER réunissent, en un même lieu, environ 80 rhumatologues libéraux ou hospitaliers, pendant deux jours. Le temps fort a lieu au cours des ateliers ; chacun de ceux-ci est animé par un membre du Conseil scientifique assisté d'un bibliographe. Chaque rhumatologue participe successivement aux trois ateliers. Schématiquement, le déroulement d'un atelier, dont la durée est de 1 heure et 45 minutes, est le suivant :

- étape 1 : brève phase d'entraînement des participants à la technique du TCS ;
- étape 2 : présentation de chaque vignette des TCS aux participants puis vote « spontané ». Il leur est demandé « *ce qu'ils font aujourd'hui dans cette situation* ». Les résultats du vote ne sont pas présentés dans l'immédiat ;
- étape 3 : présentation des « zones d'incertitude » de l'algorithme de la HAS correspondant aux TCS présentés ; puis, exposé synthétique de la revue actualisée de la littérature sur ce sujet par le bibliographe ;
- étape 4 : présentation des résultats des votes « spontanés » de l'étape 2. Puis, discussion globale des participants à propos des résultats du scrutin et des données de la littérature ;
- étape 5 : second vote éclairé par l'étape précédente où il est demandé aux participants : « *dans ces conditions, que ferez-vous demain ?* »

Les résultats du scrutin ne sont pas communiqués pendant l'atelier.

Le lendemain matin, les résultats comparatifs des deux votes successifs sont présentés aux participants, en plénière, sous forme d'histogrammes. Cette comparaison traduit et mesure donc l'impact des données actualisées de la littérature et de la discussion sur la décision « spontanée » initiale des rhumatologues confrontés aux vignettes cliniques. Une discussion d'ensemble est engagée avec les participants. Les résultats de cette phase ne seront pas présentés ici.

Une évaluation de chaque atelier et de l'ensemble des RER est faite par chaque participant. Le questionnaire d'opinion était constitué d'une échelle de Likert en 5 points allant de 1, « sans intérêt » à 5 « très utile ». Il portait sur l'intérêt que les participants éprouvaient à travailler sur les « zones d'incertitude », sur la technique du TCS, sur l'analyse de la littérature ; il les interrogeait aussi sur l'appréciation globale des RER ([Tableau 3](#)).

Tableau 1

Principes du test de concordance de script (TCS) pour l'investigation diagnostique, pour la demande d'examen complémentaire et pour la prise en charge thérapeutique.

Si vous pensiez à	Et qu'alors vous trouvez	L'effet sur votre option initiale est				
Option diagnostique	Une nouvelle information clinique	-2 Totalelement rejetée	-1 Révisée	0 Sans impact	+1 Renforcée	+2 Totalelement confirmée
Si vous pensiez demander	Et qu'alors vous trouvez	L'effet sur votre option initiale est				
Option d'examen complémentaire	Une nouvelle information clinique	-2 Totalelement rejetée	-10 Révisée	0 Sans impact	+1 Renforcée	+2 Totalelement confirmée
Si vous pensiez prescrire	Et qu'alors vous trouvez	L'effet sur votre option initiale est				
Une option thérapeutique	Une nouvelle information clinique ou résultat d'examen complémentaire	-2 Totalelement rejetée	-1 Révisée	0 Sans impact	+1 Enforcée	+2 Totalelement confirmée

Tableau 2

Exemple de vignette d'un test de concordance (TCS) concernant la prise en charge thérapeutique d'un patient atteint de polyarthrite rhumatoïde.

Madame V. a une polyarthrite rhumatoïde avec facteur rhumatoïde (+) et anti-CCP (+) depuis cinq mois. Son DAS 28 est à 2,9. Elle dit aller bien depuis qu'elle est sous méthotrexate injectable 20 mg/semaine

Si vous pensiez	Et qu'alors vous trouvez	L'effet sur votre opinion initiale est				
Maintenir le même traitement	Une nouvelle érosion sur radios mains	-2	-1	0	+1	+2

Entourez la proposition qui vous semble adéquate
 -2 : l'hypothèse est « totalement rejetée »
 -1 : l'hypothèse est « révisée »
 0 : l'information n'a « aucun impact » sur l'hypothèse
 +1 : l'hypothèse est « renforcée »
 +2 : l'hypothèse est « totalement confirmée »

Tableau 3

Satisfaction des participants aux sessions 2009 et 2010, exprimée en pourcentage de participants sur une échelle de Likert de 1 à 5.

	1 (sans intérêt)		2		3		4		5 (très utile)	
	2009	2010	2009	2010	2009	2010	2009	2010	2009	2010
L'idée de travailler sur les zones d'ombre des recommandations existantes me paraît	0	0	0	0	2,1	3,2	57,4	43,5	40,4	53,2
Je trouve la méthode du TCS	0	1,6	2,1	1,6	36,2	25,8	55,3	50,0	6,4	21,0
À l'issue de ces journées, j'estime avoir appris sur l'analyse systématique de la littérature	2,1	1,6	4,3	1,6	12,8	21,0	57,4	50,0	23,4	25,8
Je trouve le format général de ces journées	0	0	0	0	4,3	8,1	57,4	50,0	38,5	41,9
Je considère que le travail effectué lors de ces rencontres	0	1,6	2,1	6,5	27,7	6,5	48,9	54,8	21,3	30,6

TCS : test de concordance de script.

Cette enquête a été réalisée à l'issue de la session plénière finale.

2.8. Diffusion dans les régions : les points rencontre

Le contenu de ces rencontres est ensuite diffusé dans la France entière.

Ces points rencontre ont pour but de faire profiter de cette formation professionnelle, reposant sur une approche pédagogique originale, le plus grand nombre possible de rhumatologues. Ils concernent environ 600 praticiens. Le déroulement de ces réunions est calqué sur celui des ateliers des RER. Toutefois, leur durée n'excédant pas deux à trois heures, le contenu ne porte que sur les TCS et la bibliographie y afférant considérés comme les plus utiles. La sélection est faite par le Conseil scientifique en prenant en compte l'évaluation faite par les participants aux RER.

3. Résultats

Les caractéristiques des participants aux deux sessions organisées, en recourant à cette nouvelle procédure, en 2009 et 2010, sont présentées dans le [tableau 4](#).

Tableau 4

Caractéristiques des participants aux sessions 2009 et 2010 des Rencontres d'expert en rhumatologie.

	2009 (n = 72)	2010 (n = 77)
Femmes (%)	51,4	40,3
Hommes (%)	48,6	59,7
Exercice hospitalier (%)	45,8	39,0
Exercice libéral (%)	15,3	23,4
Exercice mixte ^a (%)	20,8	36,4
Exercice non précisé (%)	18,1	1,3

^a mixte : à la fois libéral et hospitalier.

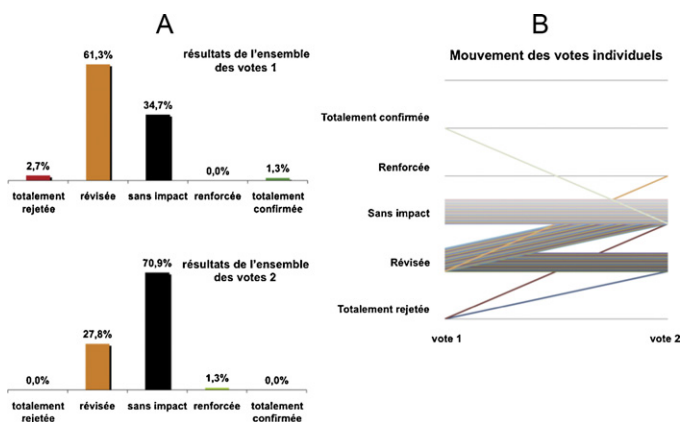


Fig. 2. Exemple à propos de l'un des tests de concordance de script (TCS) : comparaison des votes « spontanés » (votes 1) et des votes « éclairés » (votes 2) par l'exposé de la littérature actualisée et la discussion. A : évolution des votes de l'ensemble des participants ; B : évolution des votes individuels.

Un exemple de « zone d'incertitude », sélectionné par le Conseil scientifique, est exposé sur le Fig. 1. Au cours de chaque atelier, c'est-à-dire pour un domaine donné, deux à trois « zones » étaient sélectionnées, soit au total sept à huit pour l'ensemble des trois ateliers d'une session de RER donnée. Elles concernaient le diagnostic, les examens complémentaires ou la prise en charge thérapeutique de la PR.

Un exemple de TCS concernant la prise en charge thérapeutique d'un patient atteint de PR est présenté dans le tableau 2. Entre six et dix TCS étaient proposés aux participants au cours de chaque atelier, soit au total, pour les trois ateliers, entre 20 et 25 TCS. Un exemple de comparaison des scrutins observés, à propos d'un TCS, lors du 1^{er} vote (spontané) et du second (après exposé de la littérature actualisée et discussion) est présenté sur la Fig. 2.

La satisfaction des participants est présentée dans le tableau 2. Il en ressort essentiellement que « l'idée de travailler sur les zones d'ombre des recommandations... » était considérée comme utile ou très utile ($\geq 4/5$ sur échelle visuelle) par plus de 96% des participants ; si, en 2009, seuls 61,7% des participants considéraient « la méthode du TCS » comme utile ou très utile, ils étaient plus nombreux, soit 71%, à avoir cette opinion favorable en 2010.

4. Discussion

La façon dont les médecins suivent les recommandations professionnelles est un excellent moyen d'évaluer les pratiques. Paradoxalement, il existe peu ou pas de données publiées sur des outils validés pour mesurer l'impact des recommandations professionnelles au quotidien. Ce travail constitue donc une approche originale.

Jusqu'à présent, le TCS a surtout été utilisé comme instrument d'évaluation du raisonnement clinique des étudiants et des internes [16,17]. Quelques études ont rapporté l'utilisation du TCS comme outil d'apprentissage en formation continue, essentiellement en médecine générale [18]. L'expérience décrite ici est la première à montrer que cet outil pouvait également être utilisé comme outil pédagogique chez des médecins spécialistes expérimentés.

La sélection des « zones d'incertitude » dans les recommandations de la HAS dédiées à la PR s'est avérée aisée : trois tours de Delphi ont suffi au cours des sessions 2009 et 2010. Ces choix ont d'ailleurs été validés par le bon niveau de satisfaction des participants.

La technique du TCS a été développée il y a une dizaine d'années par l'équipe de B. Charlin grâce aux progrès de la psychologie cognitive [15]. Les outils écrits disponibles utilisés jusqu'alors pour

mesurer la compétence, notamment les questions à choix multiples (QCM), avaient un défaut majeur : les étudiants en médecine avaient souvent de meilleurs résultats que les médecins expérimentés. La raison en est que, probablement, l'outil utilisé – ici la QCM – ausculte essentiellement les connaissances factuelles et ne mesure donc que la capacité à résoudre des problèmes simples et bien définis. Or, en pratique, les situations sont souvent beaucoup plus complexes et mal définies. C'est l'aptitude à résoudre ces problèmes qui caractérise les médecins compétents et expérimentés.

Le TCS permet d'analyser le raisonnement clinique, en situation d'incertitude. Il a l'avantage d'évaluer l'organisation des connaissances et non pas exclusivement les connaissances factuelles. Il est fondé sur la théorie des scripts selon laquelle les professionnels progressent en réactivant leurs connaissances (scripts) qu'ils confrontent en permanence à toute nouvelle information qui leur parvient. La validité de construit du TCS, c'est-à-dire l'adéquation entre ce test et le cadre théorique dans lequel il est construit, a été vérifiée [16,19]. Le détail des qualités psychométriques du TCS a fait l'objet de plusieurs publications [20–25]. Il a été montré que les médecins expérimentés réalisaient de meilleures performances que les étudiants. Ainsi, le TCS évalue mieux l'organisation des connaissances en situation d'incertitude. Ici, contrairement aux propositions des promoteurs de la méthode, nous avons utilisé cet outil sans recourir à un panel de référence, les rhumatologues participants étant expérimentés. L'objectif n'était pas de juger leur pratique. En revanche, nous souhaitions mesurer l'impact qu'ont, sur la décision initiale (1^{er} vote « spontané »), la connaissance de la littérature actualisée et une discussion ouverte entre professionnels. Pour ce faire, un second vote (« éclairé ») a eu lieu, après exposé de la littérature et discussion. À notre connaissance, c'est la première fois qu'une telle démarche est entreprise. La mesure de cet impact fera l'objet d'une publication spécifique.

Le TCS a de nombreuses qualités [15]. Ainsi, son administration est facile sous forme papier-crayon ou sous forme informatisée ; son acceptabilité est bonne qu'il s'agisse d'étudiants ou de professionnels expérimentés. Il a toutefois quelques limites. Quelles que soient les modalités d'application, le recours au TCS nécessite un travail important de préparation, par des pédagogues entraînés à cette technique : la rédaction des vignettes est assez subtile. Ainsi, dans notre expérience, la qualité des TCS nous a semblé s'améliorer entre 2009 et 2010, les rédacteurs de ceux-ci maîtrisant mieux cette technique. C'est probablement l'une des raisons du meilleur indice de satisfaction la seconde année. Comme recommandé, les TCS ont été élaborés par un binôme du Conseil scientifique [15] ; il a, en effet été montré que construit par un seul enseignant, il s'avérait moins riche et moins pertinent et, qu'à l'opposé, réalisé par un groupe plus important il exposait à des discussions inutiles. En outre, chaque TCS nécessite une validation préalable afin d'ajuster voire de supprimer les items confus ou non pertinents. Traditionnellement, la validation est effectuée par un panel de référence ; ici, elle a été faite, en plénière, par le Conseil scientifique (étape 6 de la phase préparatoire).

Nous venons de présenter la procédure détaillée mise en œuvre au cours des RER. L'originalité de cette évaluation des pratiques professionnelles *in vitro* tient essentiellement au fait que nous avons eu recours à un outil pédagogique moderne – le TCS – combiné à l'exposé de la littérature actualisée. Cette approche s'est avérée faisable et bien acceptée par des rhumatologues libéraux et hospitaliers expérimentés, comme l'atteste le bon niveau de satisfaction. Ces résultats sont d'autant plus encourageants que la mise en œuvre de cette méthode risquait d'être rejetée par les participants. En effet, elle est quelque peu réputationnaire car elle ausculte l'incertitude et amène à s'auto-évaluer. Cette première expérience positive incite à utiliser cet outil à une plus grande échelle pour l'évaluation des pratiques professionnelles.

Déclaration d'intérêts

Les auteurs déclarent ne pas avoir de déclaration d'intérêts de conflits d'intérêts en relation avec cet article. Les RER sont organisées et soutenues par le laboratoire Abbott. Le Comité scientifique agit en toute indépendance.

Remerciements

Nous remercions S. Parisse pour la frappe du manuscrit.

Nous remercions les membres du Conseil scientifique et tous les participants aux Rencontres d'experts en rhumatologie des années 2009 et 2010.

Références

- [1] Combe B, Landewe R, Lukas C, et al. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCSIT). *Ann Rheum Dis* 2007;66:34–45.
- [2] Le Loët X, Berthelot JM, Cantagrel A, et al. Clinical practice decision tree for the choice of the first disease-modifying antirheumatic drug for very early rheumatoid arthritis: a 2004 proposal of the French Society of Rheumatology. *Ann Rheum Dis* 2006;65:45–50.
- [3] Smolen JS, Landewé R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Ann Rheum Dis* 2010;69:964–75.
- [4] American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis. *Update Arthritis Rheum* 2002;46:328–46.
- [5] Haute Autorité de santé. Recommandations professionnelles. Polyarthrite rhumatoïde. Septembre 2007. www.has-sante.fr.
- [6] Charlin B, Gagnon R, Sibert L, et al. Le test de concordance de script, un instrument d'évaluation du raisonnement clinique. *Pedagog Med* 2002;3:135–44.
- [7] Colin R. Médecine basée sur les preuves et éducation médicale. *Pedagog Med* 2001;2:69–70.
- [8] Pavy S, Constantin A, Pham T, et al. Methotrexate therapy for rheumatoid arthritis: clinical practice guidelines based on published evidence and expert opinion. *Joint Bone Spine* 2006;73:388–95.
- [9] Dalkey N, Helmer D. An experimental application of the Delphi method to the use of experts. *Manag Sci* 1963;9:458–67.
- [10] Schön DA. The reflective practitioner: how professionals think in action. New York: Basic Books; 1983.
- [11] Feltovich PJ, Barrows HS. In: Schmidt HG, De Volder ML, editors. Issues of generality in medical problem solving. Tutorials in problem-based learning: a new direction in teaching the health professions Assen (Holland): Van Gorcum; 1984. p. 128–42.
- [12] Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective on medical expertise: theory and implications. *Acad Med* 1990;65:611–21.
- [13] Charlin B, Tardif J, Boshuizen HPA. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Acad Med* 2000;75:182–90.
- [14] Charlin B, Boshuizen HPA, Custers EJFM, et al. Scripts and clinical reasoning. *Med Educ* 2007;41:1179–85.
- [15] Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. *BMC Med Inform Decis Mak* 2008;6:8–18.
- [16] Charlin B, Brailovsky CA, Brazeau-Lamontagne L, et al. Script questionnaires: their use for assessment of diagnostic knowledge in radiology. *Med Teach* 1998;20:567–71.
- [17] Sibert L, Charlin B, Corcos J, et al. Assessment of clinical reasoning competence in urology with the script concordance test: an exploratory study across two sites from different countries. *Eur Urol* 2002;41:227–33.
- [18] Labelle M, Gagnon R, Thivierge RL, et al. Formation en petits groupes sur l'ostéoporose: comparaison d'un atelier basé sur le test de concordance de scripts (TCS) et d'un atelier classique. *Pedagog Med* 2003;4:145–53.
- [19] Charlin B, Brailovsky CA, Leduc C, et al. The Diagnostic Script Questionnaire: a new tool to assess a specific dimension of clinical competence. *Adv Health Sci Ed* 1998;20:51–8.
- [20] Charlin B, Vleuten Van Der C. Standardized assessment of reasoning in context of uncertainty. The Script Concordance Test approach. *Eval Health Prof* 2004;27:304–19.
- [21] Gagnon R, Charlin B, Coletti M, et al. Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? *Med Educ* 2005;39:284–91.
- [22] Charlin B, Gagnon R, Pelletier J, et al. Assessment of clinical reasoning in the context of uncertainty: the effect of variability within the reference panel. *Med Educ* 2006;18:22–7.
- [23] Demeester A. Évaluation des processus de Raisonnement clinique des étudiants sages-femmes en fin de cursus. Mémoire de maîtrise, 13. Bobigny: Université de Paris; 2004.
- [24] Meterissian S, Zabolotny B, Gagnon R, et al. Is the script concordance test a valid instrument for assessment of intraoperative decision-making skills? *Am J Surg* 2007;193:248–51.
- [25] Gagnon R, Charlin B, Roy L, et al. The cognitive validity of the script concordance test: a processing time study. *Teach Learn Med* 2006;18:22–7.