

# Évaluation des sites web médicaux

## *Fidélité interobservateur et intraobservateur d'un outil d'évaluation*

Pierre Frémont, MD, PHD, CCMF Michel Labrecque, MD, PHD, FCMF  
France Légaré, MD, MSC, FCMF Lucie Baillargeon, MD, MSC Lucie Misson, INF

### RÉSUMÉ

**OBJECTIF** Développer un outil d'évaluation de la qualité des sites Web présentant de l'information médicale fondée sur des données probantes et en tester la fidélité de mesure.

**DEVIS** Une valeur de 60% fut attribuée aux huit critères relatifs au contenu alors qu'une valeur de 40% fut attribuée aux neuf critères relatifs à l'organisation et la présentation. Sur 10 sites sélectionnés de façon aléatoire, cinq rencontrèrent les critères d'inclusion et furent utilisés pour tester la fidélité de la grille d'évaluation par trois observateurs. Chaque site fut évalué à deux reprises à trois semaines d'intervalle par chaque observateur.

**MILIEU** Université Laval, Québec.

**PARTICIPANTS** Les trois observateurs.

**PRINCIPALES MESURES DES RÉSULTATS** Le coefficient de corrélation intraclasse (CCI) fut utilisé afin d'apprécier la fidélité de mesure de l'outil.

**RÉSULTATS** La moyenne des scores globaux des cinq sites fut: 40%, 79%, 83%, 88% et 89%. Les trois observateurs attribuèrent les quatrième et cinquième rangs aux mêmes deux sites de qualité inférieure et les trois premiers rangs aux trois autres sites. Des CCI de 0,93 à 0,97 furent obtenus pour l'évaluation globale des cinq sites par les trois observateurs. Un CCI de 0,87 fut obtenu pour l'évaluation globale des sites entre les évaluations effectuées à trois semaines d'intervalle par les trois observateurs.

**CONCLUSION** Ce nouvel outil d'évaluation des sites Web médicaux présente une excellente fidélité de mesure interobservateur et intraobservateur et permet une bonne discrimination entre des sites de qualité variable. L'utilisation simultanée de l'outil par deux observateurs et la résolution des divergences par consensus est suggérée pour optimiser la valeur des évaluations.

### ABSTRACT

**OBJECTIVE** To develop and test the reliability of a tool for rating websites that provide information on evidence-based medicine.

**DESIGN** For each site, 60% of the score was given for content (eight criteria) and 40% was given for organization and presentation (nine criteria). Five of 10 randomly selected sites met the inclusion criteria and were used by three observers to test the accuracy of the tool. Each site was rated twice by each observer, with a 3-week interval between ratings.

**SETTING** Laval University, Quebec city.

**PARTICIPANTS** Three observers.

**MAIN OUTCOME MEASURES** The intraclass correlation coefficient (ICC) was used to rate the reliability of the tool.

**RESULTS** Average overall scores for the five sites were 40%, 79%, 83%, 88%, and 89%. All three observers rated the same two sites in fourth and fifth place and gave the top three ratings to the other three sites. The overall rating of the five sites by the three observers yielded an ICC of 0.93 to 0.97. An ICC of 0.87 was obtained for the two overall ratings conducted 3 weeks apart.

**CONCLUSION** This new tool offers excellent intraobserver and interobserver measurement reliability and is an excellent means of distinguishing between medical websites of varying quality. For best results, we recommend that the tool be used simultaneously by two observers and that differences be resolved by consensus.

*Cet article a fait l'objet d'une évaluation externe.*

*This article has been peer reviewed.*

*Can Fam Physician 2001;47:2270-2278.*

**D**epuis 1993, un nombre toujours croissant de sources d'information disponibles dans Internet s'ajoute aux 12 millions d'articles et 17 000 livres médicaux publiés annuellement. On estime que plus de 100 000 sites médicaux étaient accessibles en novembre 1999<sup>1</sup>.

Ces adresses Internet offrent un large éventail de possibilités: bibliothèques virtuelles, bases de données cliniques, moteurs de recherche, réseaux d'échange, sites d'associations ou fournitures biomédicales. De plus, tout individu, groupe ou association disposant d'un ordinateur et d'un branchement Internet, peut décider de créer son propre site Web et d'y présenter des documents relatifs à la santé. Cette absence de réglementation et d'encadrement laisse donc à l'internaute recherchant de l'information médicale, la responsabilité de remettre en question et d'évaluer la qualité des informations proposées.

Comment trouver rapidement une information médicale de qualité basée sur les meilleures preuves médicales disponibles? Cette préoccupation rejoint la notion de evidence-based medicine ou médecine fondée sur les données probantes que Sackett et ses collaborateurs<sup>2</sup> définissent comme étant « [...] l'utilisation consciencieuse, explicite et judicieuse des meilleures preuves scientifiques disponibles dans le processus de décision médicale relatif à un patient ».

Afin de guider le choix du consommateur et le rassurer sur la fiabilité et la validité de l'information offerte, certains sites Internet présentent les distinctions ou honneurs qui leur sont attribués telles que: « top 5% », « sceau d'approbation », « étoiles », « médailles »<sup>3-5</sup>. Il n'est malheureusement pas toujours possible de vérifier le mécanisme d'attribution ni la méthode utilisée pour évaluer le site primé. Conscients de ce problème, plusieurs groupes à travers le monde<sup>6-11</sup> se sont donnés comme mission de guider les utilisateurs de sites Internet en santé et plusieurs d'entre eux<sup>10-12</sup> participent activement à l'élaboration de critères de qualité des sites Web médicaux. Malgré ces différentes démarches visant à promouvoir un contrôle et une réglementation plus sévère des contenus des sites Internet, aucun outil d'évaluation de la qualité des sites Web médicaux n'a fait l'objet d'une étude de fidélité interobservateur<sup>13,14</sup>.

.....  
**D<sup>rs</sup> Frémont et Labrecque** sont, respectivement, professeur associé et professeur titulaire; **D<sup>res</sup> Légaré et Baillargeon** sont professeures de clinique; et **M<sup>me</sup> Misson** est infirmière au Département de médecine familiale à l'Université Laval au Québec.

Afin de combler l'absence d'un tel instrument d'évaluation et d'en arriver à créer un répertoire critique et bilingue des sites Web médicaux, notre équipe de recherche a entrepris de développer et d'étudier la fidélité de mesure interobservateur et intraobservateur d'un outil d'évaluation de la qualité des sites Web présentant de l'information médicale fondée sur des données probantes.

## MÉTHODOLOGIE

### Développement de la grille d'analyse

Les principes généraux d'un instrument développé antérieurement par un des membres de l'équipe de recherche<sup>15</sup> furent utilisés comme point de départ. Nous avons également revu les critères proposés par les groupes qui s'étaient donné comme mission d'assurer la qualité de l'information médicale sur Internet<sup>9,10,12</sup>. Les critères jugés pertinents par consensus de tous les membres de l'équipe de recherche furent adaptés, au besoin, et intégrés afin de créer l'outil ESWM (Évaluation des Sites Web Médicaux) qui fait l'objet de la présente étude de fidélité.

Deux catégories de critères furent retenus: des critères relatifs au contenu d'une part et des critères relatifs à l'organisation et à la présentation d'autre part. De façon arbitraire, une valeur de 60% fut attribuée aux huit critères relatifs au contenu alors qu'une valeur de 40% fut attribuée aux neuf critères relatifs à l'organisation et à la présentation du site évalué (**tableau 1**). Cette répartition a pour but d'assurer, lors de l'évaluation d'un site, une prédominance de la qualité de l'information véhiculée sur l'apparence et la présentation de cette dernière. Pour chaque critère, l'évaluateur indique son niveau d'appréciation selon une échelle à trois niveaux (oui, plus ou moins, non) permettant d'attribuer au site, la totalité, la moitié ou aucun des points accordés à ce critère, respectivement. Finalement, certains critères jugés particulièrement importants ont une valeur double au sein des deux catégories. C'est le cas pour les critères « précision » et « qualité » de l'information au niveau du contenu de même que pour le critère « facilité de navigation » au niveau de l'organisation et présentation du site.

### Étude de la fidélité de mesure de la grille d'analyse

Dans un premier temps, une recherche a été lancée sur Internet à l'aide des principaux moteurs de recherche (Alta Vista, Yahoo, InfoSeek, HotBot) avec les mots-clés suivants: evidence-based medicine, données probantes, guides de pratique, practice guidelines, critical

## RECHERCHE

### Évaluation des sites web médicaux

**Tableau 1. Critères retenus pour la création de l'outil ESWM**

CRITÈRES CONTENU	60%	CRITÈRES ORGANISATION ET PRÉSENTATION	40%
But du site Web	6%	Navigation à l'intérieur du site	8%
Type d'utilisateur visé	6%	Liens efficaces entre le site et l'information originale	4%
Précision de l'information	12%	Présentation visuelle	4%
Qualité de l'information	12%	Liens avec d'autres sites pertinents	4%
Contenu applicable à la réalité canadienne	6%	Délais de téléchargements	4%
Informations mises à jour	6%	Possibilité de saisir/imprimer l'information	4%
Biais potentiel	6%	Coordonnées de l'auteur/créateur du site disponible	4%
Compétences de l'auteur	6%	Bilinguisme du site	4%
		Frais pour accéder au site	4%

**Tableau 2. Sites retenus pour tester la fidélité interobservateur et intraobservateur: Visités entre le 99/04/01 et le 99/05/31.**

Canadian Family Physician Critical Appraisal Index

<http://www.cfpc.ca/CFPC/cfpcritindex.htm>

Clinical Practice Guidelines Infobase\*

<http://www.islandnet.com/~sghollan/guidelns.htm>

National Health and Medical Research Council (Australie)

<http://www.health.gov.au/nhmrc/publicat/contents.htm>

New Zealand Health Technology Assessment

<http://nzhta.chmeds.ac.nz>

PedsCCM Evidence-Based Journal Club

[http://PedsCCM.wustl.edu/EBJournal\\_Club.html](http://PedsCCM.wustl.edu/EBJournal_Club.html)

\*Pas encore disponible.

appraisal et lecture critique de la littérature médicale. Dans un deuxième temps, les sites liens (en anglais, links) offerts par les sites obtenus furent retenus sur la base des mêmes mots-clés. Finalement, nous avons visité et retenu plusieurs sites recommandés par différentes publications médicales lues par les chercheurs. Cette démarche nous a permis d'établir une liste initiale de 150 sites pouvant contenir de l'information médicale fondée sur des données probantes.

Dans une première étape de standardisation, nous avons utilisé six de ces sites, sélectionnés de façon aléatoire, de façon à assurer une compréhension uniforme des critères d'inclusion/exclusion et des critères d'évaluation de la grille d'analyse. L'analyse de ces sites par quatre observateurs indépendants a permis d'établir, par consensus, le libellé de l'outil ESWM (figure 1) dont la fidélité de mesure fait l'objet de la présente étude.

Une fois cette étape complétée, 10 sites furent sélectionnés de manière aléatoire au sein de la liste des 150 sites initialement identifiés pour réaliser la présente étude préliminaire. De ces sites, cinq rencontrèrent les critères d'inclusion (tableau 2). Ces sites servirent à tester la fidélité interobservateur et intraobservateur de l'outil ESWM par trois observateurs indépendants. Ces trois observateurs avaient été impliqués dans le processus de développement et la procédure de standardisation de l'utilisation de l'outil ESWM.

Deux concepts furent utilisés pour évaluer la fidélité de mesure de l'outil ESWM : la fidélité de mesure intraobservateur lors d'évaluations répétées (test-retest) et la fidélité de mesure interobservateur. La fidélité de mesure interobservateur et intraobservateur fut évaluée à l'aide du coefficient de corrélation intraclasse (CCI). Le CCI consiste à évaluer la proportion de la variabilité des mesures attribuable à de vraies différences entre les différents sites évalués. Cette analyse n'évalue donc pas seulement le degré d'association entre deux mesures, comme le fait un coefficient de corrélation linéaire, mais il permet également d'apprécier le degré d'accord entre deux séries de mesures quantitatives. Le CCI permet donc une appréciation plus complète et plus sévère de la fidélité de mesure que le coefficient de corrélation linéaire. Le coefficient de corrélation de Pearson est également présenté à titre de mesure secondaire de résultats. La comparaison des cotes attribuées à chaque site par chaque observateur pris deux à deux permet de vérifier la fidélité interobservateur. Au terme d'un intervalle de 3 semaines, chacun des cinq mêmes sites fut réévalué par les trois observateurs. La comparaison des cotes attribuées lors de la première mesure à celles obtenues lors de la deuxième mesure permet d'apprécier la fidélité de mesure intraobservateur (test-retest) de la grille d'évaluation.

## RÉSULTATS

### Fidélité interobservateur

La figure 2 présente les scores attribués à chaque site par les trois observateurs lors de la première évaluation.

**Figure 1. Outil d'évaluation des sites Web médicaux (ESWM):** Version 99/03/29.

**ÉVALUATION DES SITES WEB PRÉSENTANT DES INFORMATIONS MÉDICALES FONDÉES SUR DES DONNÉES PROBANTES®**

**CRITÈRES D'ÉVALUATION**

Pierre Frémont Lucie Baillargeon Michel Labrecque France Légaré

**CRITÈRES D'INCLUSION**

Un site pertinent dans le cadre de ce projet est un site qui fournit une synthèse critique d'informations issues de la recherche médicale en vue de l'intégration en pratique clinique.

À ce titre, pour être retenu, un site doit:

- 1) Site original:
  - fournir une collection de critiques d'articles originaux
  - ou**
  - fournir un ou des guides de pratique clinique fondés sur une recension des écrits
  - ou**
  - fournir une ou des revues systématiques par critères objectifs avec ou sans méta-analyse.
- ou**
- 2) Site liens:
  - fournir une liste commentée de sites sélectionnés présentant de l'information basée sur des données probantes
- et**
- 3) → être en français et/ou en anglais

**ÉVALUATION DU CONTENU**

**C-1** Le **but** du site Web est-il clairement défini et facile à comprendre?  
(Il faudra peut-être aller à la page principale du site pour trouver cette information.)

0 = non                      3 = plus ou moins                      6 = oui

**C-2** Le **type d'utilisateur visé** est-il identifié?  
(Il faudra peut-être aller à la page principale du site pour trouver cette information.)

0 = non                      3 = plus ou moins                      6 = oui

**C-3 Précision de l'information:** la méthodologie utilisée pour l'analyse critique ou le choix de l'information est-il bien décrite? (considérant le type de site: guides de pratique, club de lecture, site organisme, site lien, revues systématiques).

0 = non                      3 = plus ou moins                      6 = oui

**C-4 Qualité de l'information:** la méthodologie utilisée pour l'analyse critique de l'information est-elle valide?  
(Même si une méthode est décrite, elle peut être subjective et non validée. Une révision par les pairs peut satisfaire ce critère.)

0 = non                      3 = plus ou moins                      6 = oui

**C-5** Le contenu est-il applicable à la **réalité médicale canadienne**?  
(Si anglais, américain, français... juger selon les auteurs des analyses critiques, la spécialité touchée.)

0 = non                      3 = plus ou moins                      6 = oui

**C-6** Les informations sont-elles **mises à jour**?  
(Sites de guides de pratique: mises à jour moins fréquente valables).

0 = aucune mise à jour indiquée

3 = une mise à jour fut effectuée il y a plus d'un mois

6 = une mise à jour fut effectuée il y a un mois ou moins

**C-7** Un commanditaire, la source de financement ou l'affiliation des auteurs du site pourrait présenter un biais potentiel relativement à l'information présentée?  
(Attention au biais plus ou moins potentiel d'une société dédiée à une discipline spécifique.)

0 = non                      3 = plus ou moins                      6 = oui

**C-8** Les compétences de l'organisme responsable et/ou des auteurs du site sont-elles pertinentes au sujet traité?

0 = non                      3 = plus ou moins                      6 = oui

*suite à la page 2274*

## RECHERCHE

.....

### Évaluation des sites web médicaux

#### ORGANISATION ET PRÉSENTATION DU SITE

**O-1** Un index, un outil de recherche, de l'hypertexte et/ou une barre d'outils est (sont-ils) disponible(s) pour faciliter la navigation au sein du site?

0 = non                      4 = plus ou moins                      8 = oui

**O-2** Existe-t-il des liens efficaces entre le site et ceux présentant l'information originale (articles de recherche pour les sites originaux et sites originaux pour les sites liens) lorsqu'elle est disponible sur Internet?

0 = non (aucune référence)

2 = plus ou moins (référence présentée sans lien hypertexte)

4 = oui (référence présentée avec lien hypertexte)

**O-3** La **présentation** est-elle attrayante et facilite-t-elle l'accès à l'information?

0 = non                      2 = plus ou moins                      4 = oui

**O-4** Existe-t-il des **liens** vers d'autres sites connexes (sites originaux ' originaux et sites liens à liens)?

0 = non                      2 = plus ou moins                      4 = oui

**O-5** Les **délais de téléchargement** perturbent-ils la navigation dans le site (délai initial et délais entre les composantes du site pour un type d'équipement donné)?

0 = non                      2 = plus ou moins                      4 = oui

**O-6** Est-il possible de **saisir et/ou d'imprimer l'information** disponible sur le site?

0 = non ou difficulté d'impression de la page visualisée

2 = impression de page visualisée

4 = contenu disponible sous forme de document électronique (p. ex., document PDF, JPEG, WORD, RTF, etc)

**O-7** La **communication** avec l'auteur/créateur du site est-elle possible?

0 = non

2 = oui, moyens conventionnels (p. ex., adresse postale/fax/tél/ courriel)

4 = oui, moyens conventionnels et forum de discussion

**O-8** Le site est-il **bilingue** (français/anglais)?

0 = non                      2 = plus ou moins                      4 = oui

**O-9** Doit-on **payer** pour accéder au site?

0 = oui                      2 = en partie                      4 = non

#### FICHE D'ÉVALUATION IDENTIFICATION ET DESCRIPTION DU SITE

Adresse WWW: \_\_\_\_\_

Nom du site: (À remplir uniquement si selon vous, ce site n'aurait pas dû être inclu.)

Inclusion:                      OUI                       NON

Raison si exclu: \_\_\_\_\_

#### ÉVALUATION DU CONTENU

**C-1** But du site Web: /6

**C-2** Type d'utilisateur visé: /6

**C-3** Précision de l'information: /12

**C-4** Qualité de l'information: /12

**C-5** Contenu applicable à la réalité canadienne: /6

**C-6** Informations mises à jour: /6

**C-7** Biais potentiel: /6

**C-8** Compétences de l'auteur: /6

**TOTAL CONTENU** /60

#### ORGANISATION ET PRÉSENTATION DU SITE

**O-1** Navigation à l'intérieur du site: /8

**O-2** Liens efficaces entre le site et l'information originale: /4

**O-3** Présentation visuelle: /4

**O-4** Liens vers d'autres sites pertinents: /4

**O-5** Délais de téléchargements: /4

**O-6** Possibilité de saisir/imprimer l'information: /4

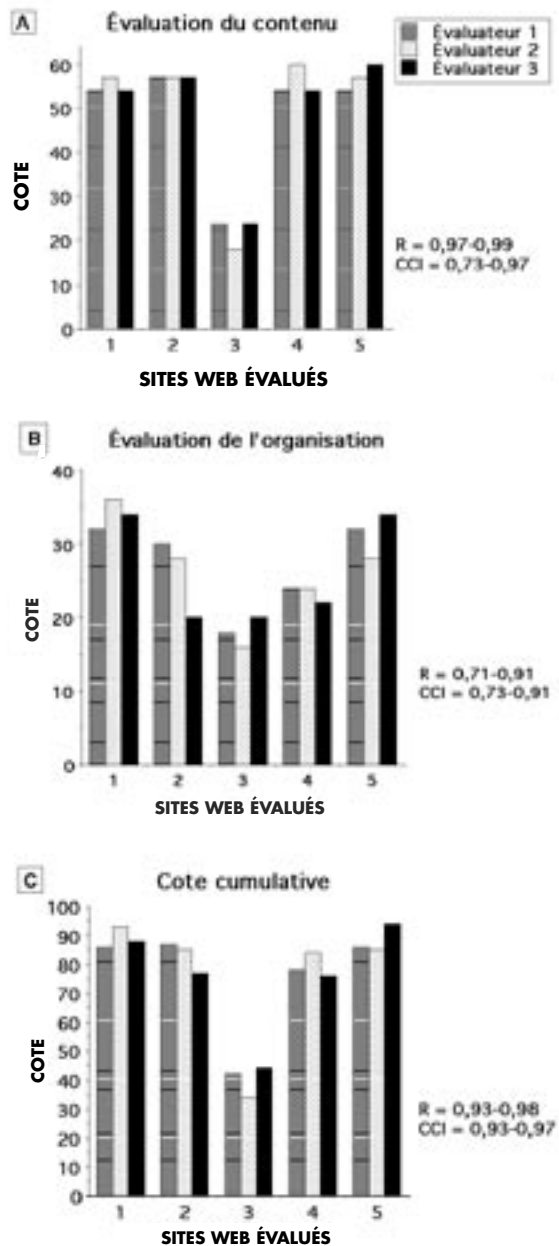
**O-7** Coordonnées de l'auteur/créateur du site disponibles: /4

**O-8** Bilinguisme du site: /4

**O-9** Frais pour accéder au site: /4

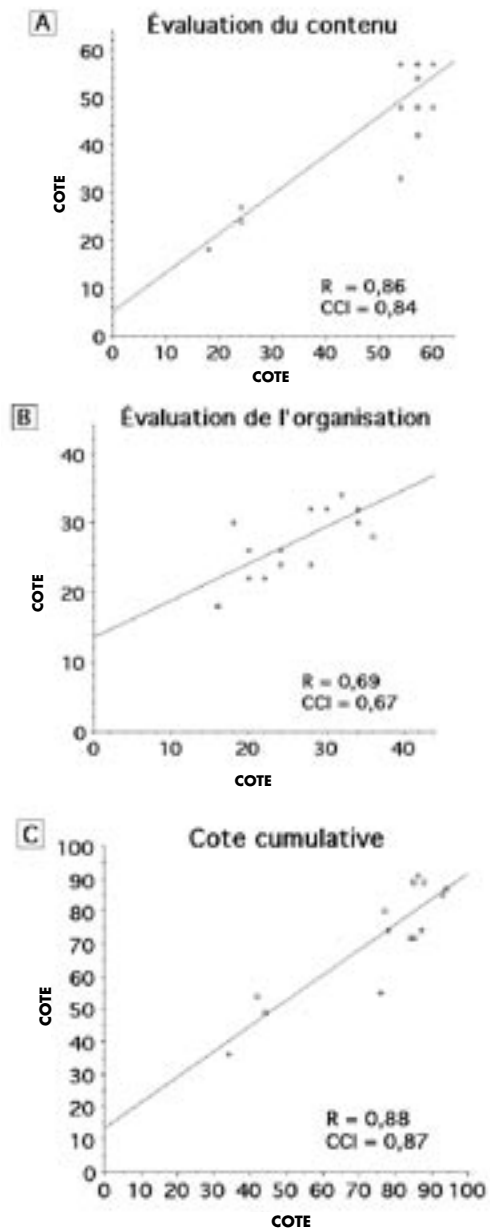
**TOTAL ORGANISATION:** /40

**Figure 2.** Étude de fidélité interobservateur de l'outil d'évaluation des sites Web médicaux ESWM: Pour l'ensemble des sites (identifiés de 1 à 5), la reproductibilité de mesure entre trois observateurs indépendants fut calculée pour A) la cote de contenu, B) la cote de présentation ainsi que pour C) la cote cumulative.



CCI—Coefficient de corrélation intraclass  
R—Coefficient de corrélation de Pearson

**Figure 3.** Étude de fidélité intraobservateur (test-retest) de l'outil d'évaluation des sites Web médicaux ESWM: La reproductibilité de mesure entre des évaluations effectuées à 3 semaines d'intervalle par trois observateurs indépendants fut calculée pour A) la cote de contenu, B) la cote de présentation ainsi que pour C) la cote globale.



CCI—Coefficient de corrélation intraclass  
R—Coefficient de corrélation de Pearson

## RECHERCHE

### Évaluation des sites web médicaux

La moyenne des scores globaux des cinq sites fut en ordre croissant de qualité: 40%, 79%, 83%, 88% et 89%. Les trois évaluateurs attribuèrent les quatrième et cinquième rangs aux mêmes deux sites de qualité inférieure et les trois premiers rangs aux trois autres sites. La comparaison des scores cumulatifs attribués aux cinq sites par chaque observateur pris deux à deux démontra des coefficients de corrélation de Pearson de 0,93 (intervalle de confiance [IC] 95%, 0,28 à 0,99;  $p = 0,02$ ), 0,94 (IC 95%, 0,34 à 0,99;  $p = 0,01$ ) et 0,98 (IC 95%, 0,74 à 0,99;  $p = 0,003$ ). L'utilisation du CCI ne change pas ces valeurs de façon notable de sorte que des valeurs de 0,93, 0,95 et 0,97 furent observées lors de la comparaison des scores cumulatifs attribués par les différentes paires d'observateurs.

La comparaison des scores attribués au contenu des sites par chaque observateur pris deux à deux démontra des coefficients de corrélation de Pearson presque parfaits (0,97 à 0,99;  $p = 0,004$  à 0,001) et les mêmes comparaisons effectuées à l'aide du CCI démontrèrent des valeurs de 0,73, 0,96 et 0,97. L'évaluation de la fidélité interobservateur pour l'évaluation de l'organisation et de la présentation des sites démontra pour sa part des coefficients de corrélation de Pearson de 0,71 ( $p > 0,05$ ), 0,71 ( $p > 0,05$ ) et 0,91 ( $p = 0,03$ ) et des CCIs de 0,73, 0,75 et 0,91.

#### Fidélité intraobservateur (test-retest)

Un coefficient de corrélation de 0,88 (IC 95%, 0,67 à 0,96,  $p < 0,0001$ ) fut obtenu pour la comparaison du score cumulatif entre les évaluations des mêmes sites effectuées à 3 semaines d'intervalle par les trois évaluateurs. La valeur correspondante pour le contenu et l'organisation des sites fut 0,86 (IC 95%, 0,62 à 0,95,  $p < 0,0001$ ) et 0,69 (IC 95%, 0,28 à 0,89,  $p = 0,004$ ), respectivement (**figure 3**). L'évaluation de la fidélité intraobservateur à l'aide du CCI démontra des valeurs de 0,87 pour le score cumulatif, 0,84 pour le contenu et 0,67 pour la présentation et l'organisation.

## DISCUSSION

Nos résultats démontrent que, dans les conditions de la présente étude, l'outil ESWM présente une excellente fidélité de mesure et permet une bonne discrimination entre des sites de qualité variable. Cette fidélité de mesure est principalement attribuable à l'évaluation du contenu et à la prédominance de cette composante dans le score cumulatif. Cette situation respecte notre préoccupation initiale qui était d'éviter qu'une présentation optimale puisse masquer un contenu insignifiant. La qualité de ces résultats fut aussi vraisemblablement favorisée par l'excellent niveau de standardisation préalable de

l'outil ESWM. En effet, les trois évaluateurs impliqués avaient non seulement participé à l'étape de standardisation décrite dans cette étude mais avaient également contribué au développement initial de l'outil. À l'inverse, il faut présumer qu'une telle standardisation pourrait s'avérer nécessaire avant de supposer qu'un tel outil puisse être utilisé avec une fidélité de mesure comparable dans un autre contexte. Lors d'études ultérieures, les conditions de standardisation préalables devront donc être précisées.

Lorsque l'on considère strictement l'évaluation de l'organisation et de la présentation des sites, une fidélité de mesure beaucoup moins optimale est cependant observée. Le degré moindre de corrélation observé dans l'appréciation de l'organisation des sites pourrait résulter en partie, des caractéristiques des postes et des liens informatiques utilisés par les différents observateurs (p. ex. modem câble vs modem téléphonique).

Au-delà des résultats satisfaisants obtenus en regard de l'objectif principal de ce projet, nous avons observé que les critères retenus pour l'outil ESWM faisaient consensus au sein des éléments d'évaluation sélectionnés par plusieurs groupes<sup>8-13</sup>. D'autre part, une étude des documents produits par ces organisations a mis en évidence d'importantes différences au niveau de l'élaboration du contenu de chacun des critères. Pour certains, un simple énoncé de principe qualifie chaque composante de l'évaluation<sup>9</sup>, d'autres présentent des critères explicitement définis<sup>8</sup> et quelques-uns d'entre eux proposent des documents d'évaluation complets et bien structurés<sup>10,12</sup>. À l'instar de Jadad et Gagliardi<sup>14</sup>, nous n'avons cependant retrouvé aucune étude de validation des grilles proposées susceptible d'être utilisée comme mesure étalon.

À la lumière des résultats de la présente étude, l'outil ESWM pourrait devenir un élément de solution dans une démarche globale visant à privilégier les sites médicaux de qualité dans l'Internet. Il faut cependant rappeler qu'en l'absence de standard de référence, tout choix et toute pondération de critères d'une grille demeure arbitraire. Malgré la rigueur du processus que nous avons suivi pour développer l'outil et qui lui confère une haute validité apparente (face validity), les résultats de l'étude de fidélité sont recevables uniquement si on accepte les critères d'évaluation utilisés (**figure 1**).

La cotation des cinq sites retenus pour l'étude de fidélité interobservateur et intraobservateur à l'aide de l'outil ESWM a également permis de mettre en lumière certaines difficultés pouvant être rencontrées lors de l'évaluation de sites de formats différents.

**Tableau 3. Précisions apportées aux critères de précision et qualité de l'information à la suite de l'étude de fidélité: A) C-3 Précision de l'information: la méthodologie utilisée pour l'analyse critique ou le choix de l'information est-il bien décrit? B) C-4 Qualité de l'information: la méthodologie utilisée pour l'analyse critique de l'information est-elle valide? (Même si une méthode est décrite, elle peut être subjective et non validée.)**

**A****GUIDES DE PRATIQUE**

1. Recension systématique des données de littérature: **6 pts**  
Recension des données de la littérature: **3 pts**
2. Recommandation appuyée sur le niveau de preuves scientifiques (évidence): **6 pts**  
Non: **0 pts**

**REVUES SYSTÉMATIQUES**

1. Recension des données de la littérature: Complète (plusieurs banques de données): **6 pts**  
Partielle (p.ex., MEDLINE seul): **3 pts**  
Non définie: **0 pts**
2. Critères d'admissibilité des études: Explicites et sélection indépendante par 2 réviseurs: **6 pts**  
Explicites seulement: **3 pts**  
Non: **0 pts**

**SITES LIENS**

1. Processus de sélection des sites explicite: **6 pts**  
Processus de sélection des sites implicite: **3 pts**
2. Description structurée: **6 pts**  
Description non structurée: **3 pts**

**RÉSUMÉS CRITIQUES D'ARTICLES ORIGINAUX**

1. Processus de sélection des articles explicite (question clinique, filtre méthodologique, choix journaux): **3 pts**  
Non: **0 pts**
2. Résumé: **3 pts**  
Structuré (population, intervention, résultats): **3 pts**  
Commentaire critique: **3 pts**

**B****GUIDES DE PRATIQUE**

1. Évaluation méthodologique des études:  
Par critères objectifs: **6 pts**  
Sans critères objectifs: **3 pts**  
Aucune évaluation méthodologique: **0 pts**
2. Révision par les pairs: **3 pts**
3. Validation en clinique: **3 pts**

**REVUES SYSTÉMATIQUES**

1. Évaluation méthodologique des études:  
À l'aide d'une grille par critères objectifs: **6 pts**  
Sans grille par critères objectifs: **3 pts**  
Aucune: **0 pts**
2. Méthode explicite pour combiner les résultats:  
Oui: **6 pts** ± **3 pts** Non: **0 pts**

**SITES LIENS**

1. Évaluation des sites à l'aide de critères explicites: **6 pts**  
Évaluation des sites à l'aide de critères implicites: **3 pts**
2. Processus de révision de la méta-information par plus d'un évaluateur: **6 pts**  
Non: **0 pts**

**RÉSUMÉS CRITIQUES D'ARTICLES ORIGINAUX**

1. Évaluation de la méthodologie:  
À l'aide d'une grille par critères objectifs: **6 pts**  
Sans grille par critères objectifs: **3 pts**  
Aucune: **0 pts**
2. Commentaire sur l'impact de l'étude:  
Importance des résultats: **3 pts**  
Applicabilité dans la pratique: **3 pts**

En effet, les sites médicaux s'adressant aux professionnels de la santé présentent des contenus de type EBM variés qui, à la lumière de nos recherches, peuvent être divisés en quatre grandes catégories: les sites présentant des guides de pratique, ceux présentant des revues systématiques, les clubs de lecture et les sites liens ou sites catalogues. De même, il est possible pour un site d'offrir un contenu répondant à plus d'une catégorie. Dans l'optique d'une utilisation de l'outil ESWM pour la création d'un répertoire des ressources médicales fondées sur les données probantes, les critères évaluant respectivement la précision et la qualité de l'information (C-3, C-4) ont été précisés afin de s'assurer que le processus d'élaboration des différents types de contenus (guides de pratique, revues de la littérature, critiques d'articles originaux) respecte les critères d'évaluation

méthodologique reconnus de la médecine fondée sur les données probantes<sup>16</sup> (tableau 3). L'application de ces critères devrait améliorer la fidélité de mesure de l'outil lors d'études ultérieures.

Finalement, afin d'optimiser la validité des évaluations, nous suggérons l'utilisation simultanée de l'outil par deux observateurs indépendants et la résolution des divergences par consensus. Concrètement, lorsque nous utilisons la grille ESWB, nous recourons maintenant à une résolution des divergences par consensus dès qu'une différence de plus de 10% est observée au niveau de l'une ou l'autre des composantes du score (contenu, organisation/présentation).

Il faut également mettre en perspective la valeur des résultats obtenus en regard de la taille de l'échantillon utilisé. Suite à l'élimination, lors de l'application des



## RECHERCHE

### Évaluation des sites web médicaux

critères d'inclusion/exclusion, de cinq des 10 sites initialement sélectionnés de façon aléatoire à partir des 150 sites identifiés, seulement cinq sites ont fait l'objet de la présente étude de fidélité. Malgré la taille réduite de cet échantillon, les résultats obtenus sont prometteurs et justifient la poursuite de l'évaluation de ce type de démarche visant à identifier l'information médicale de qualité dans l'Internet.

Dans les conditions de standardisation de la présente étude, l'outil ESWM offre une excellente fidélité de mesure interobservateur et intraobservateur et permet d'obtenir une bonne discrimination entre des sites Web de qualité variable. L'utilisation de critères de qualité et de précision de l'information spécifiques à chaque type de site évalué et l'utilisation simultanée de l'outil par deux observateurs indépendants ayant recours au consensus pour résoudre les divergences d'évaluation supérieures à 10% ne pourrait qu'accroître cette fidélité de mesure.

L'outil permettra d'analyser les sites médicaux présents dans l'Internet, de retenir ceux offrant une information médicale de qualité basée sur les meilleures preuves médicales disponibles et de les présenter aux internautes sous forme d'un répertoire critique bilingue des sites Web médicaux intégré à un site Internet entièrement dédié à la médecine fondée sur les données scientifiques. Au moment de la révision finale de cet article, ce site Web pouvait être consulté à l'adresse Internet suivante: <http://www.medecine.quebec.qc.ca>. ❖

### Remerciements

*Cette étude fut réalisée à l'aide d'une subvention de la Régie régionale de la santé et des services sociaux de Québec et de Santé Canada.*

### Contributions des auteurs

*Tous les auteurs ont participé à la conception et à la réalisation du projet, ont participé à la collecte des données, et ont approuvé la version finale du manuscrit. D<sup>r</sup> Frémont et M<sup>me</sup> Misson ont rédigé les différents ébauches du manuscrit qui ont été révisées par D<sup>rs</sup> Labrecque, Légaré, et Baillargeon. Leurs suggestions et commentaires ont été intégrés à la version finale du manuscrit.*

### Intérêts concurrents

*Rien déclaré*

**Correspondance à:** Pierre Frémont, Unité médecine familiale, CHUQ (Pavillon CHUL), 2705, boul. Laurier, Ste-Foy QC G1V 3A7; courriel [Pierre.Fremont@crchul.ulaval.ca](mailto:Pierre.Fremont@crchul.ulaval.ca)

### Références

1. Eysenbach G, Dieppen TL. Shopping around the Internet today and tomorrow: towards the millennium of cybermedicine. *BMJ* 1999;319:1294-8.

### Points de repère du rédacteur

- Cet article décrit un outil d'évaluation de la qualité des sites Web présentant de l'information médicale fondée sur des données probantes.
- L'outil a été appliqué par trois évaluateurs à cinq sites Web à deux reprises à 3 semaines d'intervalle. La fidélité intra- et inter-observateur est excellente et l'outil permet de discriminer entre des sites de qualité variable.
- L'utilisation simultanée de l'outil par deux observateurs et la résolution des divergences par consensus est suggérée pour optimiser la valeur des évaluations.

### Editor's key points

- This article describes a tool for rating the quality of websites that provide information on evidence-based medicine.
- The tool was used by three evaluators on five websites at two different times 3 weeks apart. Interobserver and intraobserver measurement reliability was excellent; the tool can distinguish between medical websites of varying quality.
- For best results, the tool should be used by two observers simultaneously, and differences should be resolved by consensus.

2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence-based medicine: what it is and what is isn't. *BMJ* 1996;312:71-2.
3. Award Site! Rating Level Information (2001). Award Sites. Ratings, Recognition, Resources ... and More. [site web]. Disponible à <http://www.awardsites.com/levels/rating-info.htm>. Page consultée le 25 janvier 2001.
4. Super! Website Awards: Awards & Criteria (2001). Website Awards [site web]. Disponible à <http://websiteawards.xe.net/superb.htm>. Page consultée le 25 janvier 2001.
5. Médaille d'Or Award: Selection Criteria (2001). Médaille d'Or Award [site web]. Disponible à <http://www.arachnid.co.uk/award/select.html>. Page consultée le 25 janvier 2001.
6. eHealth Code of Ethics (2000). Internet Healthcare Coalition [site web]. Disponible à <http://www.ihealthcoalition.org/ethics/code0524.pdf>. Page consultée le 25 janvier 2001.
7. Introduction to the British Healthcare Internet Association (2000). British Healthcare Internet Association [site web]. Disponible à <http://www.bhia.org/about>. Page consultée le 25 janvier 2001.
8. OMNI Evaluation Guidelines (2001). Organising Medical Network Association [site web]. Disponible à <http://www.omni.ac.uk/about/guide.html>. Page consultée le 25 janvier 2001.
9. HONcode: Principles (1997). Health on the Net Foundation [site web]. Disponible à <http://www.hon.ch/HONcode/Conduct.html>. Page consultée le 25 janvier 2001.
10. Net Scoring: Critères de qualité de l'information de santé sur l'Internet (2001). Centrale Santé Paris [site web]. Disponible à <http://www.chu-rouen.fr/netscoring/>. Page consultée le 25 janvier 2001.
11. Winker MA, Flanagan A, Chi-Lum B, White J, Andrews K, Kennett RL, et al. Guidelines for medical and health information sites on the Internet. *JAMA* 2000;283:1600-6.
12. Criteria for Assessing the Quality of Health Information on the Internet (1999). Health Summit Working Group [site web]. Disponible à <http://hitweb.mitretsk.org/docs/policy.pdf>. Page consultée le 25 janvier 2001.
13. Kim P, Eng TR, Deering MJ, Maxfield A. Published criteria for evaluating health related web sites: review. *BMJ* 1999;318:647-9.
14. Jadad AR, Gagliardi A. Rating health information on the Internet. *JAMA* 1998;279:611-61.
15. Ellis C, Frémont P, Quigg MT. Rating medical Web sites for family physicians. Presented at the 18th annual workshop of the Section of Teachers, College of Family Physicians of Canada; Ottawa, Ont; October 1996.
16. Sackett DL, Strauss SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. How to practise and teach EBM. 2nd ed. London, Engl: Churchill Livingstone; 2000.