



HAL
open science

Interpersonal skills in medical consultation: assessment and improvement for the quality of care

Alexandre Bellier

► **To cite this version:**

Alexandre Bellier. Interpersonal skills in medical consultation: assessment and improvement for the quality of care. Human health and pathology. Université Grenoble Alpes [2020-..], 2022. English. NNT: 2022GRALS012 . tel-03917156

HAL Id: tel-03917156

<https://theses.hal.science/tel-03917156v1>

Submitted on 1 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : MBS - Modèles, méthodes et algorithmes en biologie, santé et environnement

Arrêté ministériel : 25 mai 2016

Présentée par

Alexandre BELLIER

Thèse dirigée par **José LABARERE**
et codirigée par **Philippe CHAFFANJON**, UGA

préparée au sein du **Laboratoire Translational Innovation in Medicine and Complexity**
dans l'**École Doctorale Ingénierie pour la Santé la Cognition et l'Environnement**

Les compétences relationnelles en consultation médicale : évaluation et développement pour la qualité des soins

Interpersonal skills in medical consultation: assessment and improvement for the quality of care

Thèse soutenue publiquement le **31 mai 2022**,
devant le jury composé de :

Monsieur JOSE LABARERE

Professeur des Univ. - Praticien hosp., UNIVERSITE GRENOBLE ALPES
, Directeur de thèse

Monsieur PHILIPPE CHAFFANJON

Professeur des Univ. - Praticien hosp., UNIVERSITE GRENOBLE ALPES, Co-directeur de thèse

Monsieur MARC BRAUN

Professeur des Univ. - Praticien hosp., UNIVERSITE DE LORRAINE,
Rapporteur

Monsieur LAURENT BOYER

Professeur des Univ. - Praticien hosp., AIX-MARSEILLE UNIVERSITE,
Rapporteur

Madame MARIE-CLAUDE AUDETAT

Professeur associé, Université de Genève, Examinatrice

Madame CAROLE CANALI SCHWEBEL

Professeur des Univ. - Praticien hosp., UNIVERSITE GRENOBLE ALPES, Présidente

SOMMAIRE

RESUME.....	5
ABSTRACT.....	6
REMERCIEMENTS.....	7
1. INTRODUCTION	9
1.1. <i>Contexte</i>	9
1.1.1. Compétences relationnelles : clarification conceptuelle	9
1.1.2. Association des compétences relationnelles avec la qualité des soins.....	11
1.1.3. Importance des compétences relationnelles médecin-patient.....	12
1.2. <i>Rationnel</i>	13
1.2.1. Évaluation.....	13
1.2.2. Facteurs associés avec la mesure produite	14
1.2.3. Développement des compétences relationnelles	16
1.3. <i>Objectifs</i>	16
2. CHAPITRE 1 : ÉVALUATION DES COMPETENCES RELATIONNELLES MEDECIN-PATIENT	19
2.1. Partie 1) Revue systématique de la littérature des échelles d'évaluation des compétences relationnelles en consultation médicale	19
Article soumis — Content and psychometrics properties of standardized instruments measuring physician interpersonal skills in consultation: a Systematic Review	19
2.2. Partie 2) Adaptation transculturelle et validation des propriétés psychométriques de l'échelle 4-Habits Coding Scheme (4-HCS)	42
Article publié — Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills	42
3. CHAPITRE 2 : DETERMINANTS DES COMPETENCES RELATIONNELLES EN CONSULTATION MEDICALE	61
3.1. Partie 3) Compétences relationnelles en médecine générale ambulatoire	61
Article soumis — Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en médecine générale ambulatoire.....	61
3.2. Partie 4) Compétences relationnelles en maïeutique	76
Article publié — Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en maïeutique.....	76
3.3. Partie 5) Facteurs associés aux compétences relationnelles chez les étudiants en médecine en consultations simulées	83
Article publié — Student characteristics associated with interpersonal skills in medical consultations.....	83
3.4. Partie 6) Impact de la durée de consultation sur les compétences relationnelles des médecins	94
Impact of consultation length on interpersonal skills between medical practitioner and patient: a cross-sectional study.....	94
4. CHAPITRE 3 : DEVELOPPEMENT DES COMPETENCES RELATIONNELLES PAR DES FORMATIONS DEDIEES EN EDUCATION MEDICALE	116
4.1. Partie 7) Évaluation des dispositifs d'amélioration des compétences relationnelles lors de la formation initiale	116

Article publié — Impact of physician–patient relationship training on medical students’ interpersonal skills during simulated medical consultations. A cross-sectional study	116
4.2. Partie 8) Protocole d’un essai contrôlé randomisé pour l’évaluation d’un programme pédagogique dédié aux compétences relationnelles en formation continue 126	
Article publié — Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): Protocol for a randomised controlled trial	126
4.3. Partie 9) Résultats préliminaires de l’étude EPECREM	136
5. SYNTHESE ET PERSPECTIVES	142
5.1. <i>Évaluation des compétences relationnelles médecin-patient</i>	142
5.2. <i>Déterminants des compétences relationnelles en consultation médicale</i>	146
5.3. <i>Développement des compétences relationnelles par des programmes pédagogiques dédiés</i>	149
5.4. <i>Perspectives</i>	152
5.5. <i>Conclusion</i>	154
6. REFERENCES BIBLIOGRAPHIQUES	156
7. VALORISATION DES TRAVAUX DE THESE	165
7.1. Article 1 : Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J. Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills. PLoS ONE. 2020;15(4):e0230672.	165
7.2. Article 2 : Blanc G, Bellier A, Di Marco L. Validation du système de codage des quatre habitudes pour l’évaluation des compétences relationnelles en maïeutique. Sages-Femmes. 2020 Jul 1;19(4):50–4.	165
7.3. Article 3 : Bellier A, Chaffanjon P, Morand P, Palombi O, Francois P, Labarère J. Student characteristics associated with interpersonal skills in medical consultations. BMC Medical Education. 3 mai 2022;22(1):338.	165
7.4. Article 4 : Bosméan L, Chaffanjon P, Bellier A. Impact of physician-patient relationship training on medical students’ interpersonal skills during simulated medical consultations: a cross-sectional study. BMC Med Educ. 2022 Feb 22;22(1):117.	165
7.5. Article 5 : Bellier A, Labarère J, Putkaradze Z, Cavalie G, Carras S, Pelen F, et al. Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): protocol for a randomised controlled trial. BMJ Open. 2022 Feb 15;12(2):e051600.	165
8. ANNEXES DES ARTICLES	166
8.1. <i>Partie 1) Appendix 1</i> :	166
8.2. <i>Partie 1) Appendix 2</i> :	166
8.3. <i>Partie 1) Appendix 3</i> :	167
8.4. <i>Partie 2) Appendix 1: Cross-cultural adaptation of the 4-HCS scale into French</i> 167	
8.5. <i>Partie 2) Appendix 2: Scree plot of the 23 items of the 4-HCS scale</i>	170
8.6. <i>Partie 5) Appendix 1</i>	171
8.7. <i>Partie 5) Appendix 2</i>	171
8.6. <i>Partie 5) Appendix 3</i>	172
8.7. <i>Partie 6) Appendix 1</i>	173
8.8. <i>Partie 6) Appendix 2</i>	174
8.9. <i>Partie 6) Appendix 3</i>	175

Résumé

Les compétences relationnelles médecin-patient, qui comprennent la communication, l'empathie et la structuration de l'entretien centré sur le patient, font partie des compétences fondamentales d'un médecin. Elles constituent un élément clé d'une consultation médicale efficace et elles ont un fort impact sur la qualité des soins.

Le premier objectif était l'évaluation de ces compétences non-techniques, jugées souvent comme subjectives. Nous avons identifié l'échelle "Four Habits Coding Scheme" (4-HCS), un instrument standardisé conçu pour une hétéro-évaluation des compétences relationnelles des médecins, sur la base de consultations vidéo-enregistrées. Nous avons réalisé l'adaptation transculturelle en français de cette échelle, puis évalué ses propriétés psychométriques. La cohérence interne était satisfaisante mais l'échelle nécessitait l'utilisation de deux évaluateurs indépendants pour atteindre une fiabilité satisfaisante.

Par ailleurs, le deuxième objectif était de comprendre les déterminants des compétences relationnelles pour pouvoir développer des formations dédiées. Dans une étude observationnelle, les étudiants en médecine avec un niveau plus élevé de compétences relationnelles en simulation étaient ainsi plus susceptibles d'être des femmes, d'avoir effectué un stage clinique international (programme d'échange ERASMUS) ou en laboratoire de recherche. La durée de la consultation était également associée aux compétences relationnelles. Enfin, de nombreuses organisations ont mis en place des programmes de formation structurés pour améliorer les compétences relationnelles des médecins, avec un niveau de preuve limité de leur efficacité. Pour répondre à ce dernier objectif, nous avons donc évalué l'efficacité d'un programme standardisé et multifacettes de développement des compétences relationnelles pour les médecins hospitaliers, ainsi que des dispositifs existants pour les étudiants en médecine. Nous avons montré l'intérêt des formations dédiées, en particulier lorsqu'elles sont répétées et même si elles sont de courte durée. Notre essai contrôlé randomisé suggère un impact sur le patient par une amélioration de l'alliance thérapeutique, mais aussi sur le médecin pour son accomplissement personnel.

Mots-clés : Compétences relationnelles; Communication ; Relation médecin-patient ; Qualité des soins ; Pédagogie ; Simulation non-technique ; Evaluation

Abstract

Physician-patient interpersonal skills, which include communication, empathy and patient-centered interview structuring, are among the core skills of a physician. They are a key component of an effective medical consultation and have a strong impact on the quality of care. The first objective was the evaluation of these non-technical skills, which are often considered subjective. We identified the Four Habits Coding Scheme (4-HCS), a standardized instrument designed for independent evaluation of physicians' interpersonal skills, based on video-recorded consultations. We carried out a cross-cultural adaptation of this scale in French, and then evaluated its psychometric attributes. The internal consistency was satisfactory but the scale required the use of two independent raters to achieve satisfactory reliability.

The second objective was to elucidate the determinants of interpersonal skills in order to develop dedicated medical education program. As part of an observational study, medical students with higher levels of interpersonal skills in simulation were more likely to be female, to have completed an international clinical placement (as part of ERASMUS exchange program) or to have worked in a research laboratory. The length of the consultation was also associated with interpersonal skills.

Finally, many organizations have implemented structured training programs to improve physicians' interpersonal skills, with limited level of evidence on their effectiveness. To address the latter objective, we therefore evaluated the effectiveness of a standardized, multifaceted interpersonal skills development program for hospital-based physicians, as well as existing programs for medical students. We showed the value of dedicated training, especially when repeated and even if short in duration. Our randomized controlled trial suggested an impact on patient outcomes including therapeutic alliance, but also on the physician outcomes including personal achievement.

Keywords: Interpersonal skills; Communication; Physician-patient relationship; Quality of care; Medical Education; Non-technical simulation; Evaluation

Remerciements

A Mme le Professeur Marie-Claude Audétat, je vous remercie d'avoir accepté de participer à ce jury en tant qu'experte reconnue en éducation médicale.

A M. le Professeur Laurent Boyer, merci d'avoir accepté d'être rapporteur de ce travail. Vos travaux de recherche sont un modèle.

A M. le Professeur Marc Braun, merci de me faire l'honneur de rapporter ce travail de thèse. Votre investissement pour la pédagogie est une source d'inspiration quotidienne.

A M. le Professeur Philippe Chaffanjon, merci pour la liberté accordée depuis de nombreuses années, pour le soutien sans faille et pour avoir initié l'idée de ces travaux.

A M. le Professeur José Labarère, merci pour la grande confiance accordée et pour le plaisir de travailler ensemble, avec un regard à la fois exigeant et bienveillant.

A Mme le Professeur Carole Schwebel, merci d'avoir accepté de juger ce travail et merci pour votre large contribution au dynamisme de notre faculté.

Merci à tous les collègues qui ont participé à ces travaux (Lucie, Gwendoline, Romain, Lara, Joey, Zaza, Adeline, Ariane), merci au Département des Arts du Spectacle, en particulier à Pauline Bouchet, et merci à tous les participants étudiants ou professionnels.

Je remercie également tous ceux qui travaillent avec moi au SEEM, au CIC-P et au LADAF.

A Noémie, pour les petits bonheurs du quotidien.

A ma famille, pour m'avoir permis d'en arriver là et pour tout ce qu'on partage ensemble.

A mes amis, pour tous ces bons moments indispensables passés en votre compagnie.

A tous les passionnés d'éducation médicale qui liront ce travail avec, je l'espère, un certain intérêt.

INTRODUCTION

1. INTRODUCTION

1.1. Contexte

1.1.1. Compétences relationnelles : clarification conceptuelle

Les compétences interpersonnelles en médecine désignent des compétences relationnelles. Nous distinguerons les compétences relationnelles au sein de la relation médecin-patient, des compétences interprofessionnelles du médecin au sein d'une équipe soignante. Bien entendu, les compétences relationnelles appliquées au patient ou à l'équipe de soin sont liées et le développement de l'une affectera l'évolution de l'autre (Watters et al. 2015; Nelson et al. 2017; Berger-Estilita et al. 2020). Toutefois, nous traiterons dans ce travail uniquement des compétences relationnelles médecin-patient, généralement dans le cadre d'une consultation médicale.

Les compétences relationnelles peuvent être définies comme les compétences liées à l'interaction réciproque entre deux ou plusieurs personnes, en particulier le médecin et le patient pour la relation médecin-patient (Tableau 1). Elles font partie des compétences sociales, c'est-à-dire l'ensemble des capacités personnelles requises pour interagir et communiquer avec succès avec les autres, tant verbalement que non verbalement par les gestes, le langage corporel ou l'apparence personnelle.

Les compétences relationnelles sont souvent assimilées à tort aux compétences en communication (Duffy et al. 2004). Or, la communication tend à être définie comme l'échange ou la transmission d'idées, d'attitudes ou de croyances entre des individus. La communication ne contient généralement pas la manière, le moment et la personne à qui les paroles sont transmises (Tableau 1). Ces derniers éléments de la communication portent souvent le contenu émotionnel du discours. L'ensemble de ces éléments sont généralement inclus dans les compétences relationnelles (Dyche 2007). Ainsi, les compétences relationnelles ne se limitent pas à la communication verbale et non verbale. Il faut distinguer l'approche de compétence communicative basée sur des habiletés principalement techniques et l'approche de relation humaine entre deux individus conscients (Zoppi and Epstein 2002). Les habiletés de communication peuvent facilement être objectivées par l'observation, tandis que l'approche relationnelle dépend du patient, de son degré de satisfaction et de la raison pour laquelle le patient cherche à se faire soigner (Egener and Cole-Kelly 2004).

Les compétences relationnelles en consultation médicale comprennent donc à la fois la communication, mais aussi la structuration de l'entretien dans une relation centrée sur le patient,

l'intelligence émotionnelle et l'empathie (Krupat et al. 2006). L'intelligence émotionnelle correspond à la capacité de comprendre et de gérer les émotions, d'utiliser la connaissance des émotions pour améliorer la réflexion et traiter efficacement les tâches. L'empathie est une composante de l'intelligence émotionnelle. Elle est définie quant à elle par la conscience objective et perspicace qu'à un individu des sentiments et du comportement d'une autre personne (Tableau 1).

Les compétences relationnelles en consultation médicale incluent la capacité à recueillir les informations cliniquement pertinentes, à comprendre les attentes du patient, à motiver le patient et à fournir des conseils thérapeutiques appropriés, tenant compte du patient, le tout dans une relation bienveillante et constructive (Dyche 2007; Ha and Longnecker 2010).

Compétences sociales	L'ensemble des capacités personnelles requises pour interagir et communiquer avec succès avec les autres, tant verbalement que non verbalement par les gestes, le langage corporel et l'apparence personnelle.
Compétences relationnelles	Les compétences liées à l'interaction réciproque de deux ou plusieurs personnes.
Compétences relationnelles médecin-patient	Les compétences liées à l'interaction entre le médecin et le patient.
Communication	L'échange ou la transmission d'idées, d'attitudes ou de croyances entre des individus ou des groupes.
Communication en santé	L'étude et l'utilisation de stratégies de communication pour informer et influencer les décisions individuelles et communautaires qui améliorent la santé.
Intelligence émotionnelle	La capacité de comprendre et de gérer les émotions et d'utiliser les connaissances émotionnelles pour améliorer la réflexion et traiter efficacement les tâches. L'intelligence émotionnelle est une mesure de la capacité d'une personne à socialiser ou à entrer en relation avec les autres.
Empathie	La conscience objective et perspicace qu'a un individu des sentiments et du comportement d'une autre personne. Elle doit être distinguée de la sympathie, qui est généralement non objective et non critique. Elle inclut la sollicitude, qui est la manifestation d'une conscience et d'un souci du bien d'autrui.

Tableau 1. Arborisation et définitions des concepts utilisés

(d'après thésaurus MeSH, National Center for Biotechnology Information, U.S. National Library of Medicine)

1.1.2. Association des compétences relationnelles avec la qualité des soins

Les compétences relationnelles médecin-patient sont des compétences cliniques fondamentales qui facilitent l'établissement d'une relation de confiance entre le médecin et le patient, amenant à une alliance thérapeutique (Chichirez and Purcărea 2018) (Tableau 2). Les compétences relationnelles sont ainsi un déterminant clé d'une consultation médicale efficace et plus important encore, elles représentent les qualités les plus attendues par le patient, devant les compétences techniques (Simpson et al. 2011) ou à niveau équivalent par rapport aux compétences techniques et scientifiques (Murakami et al. 2010).

Les compétences relationnelles sont associées positivement à la satisfaction du patient lors de la consultation (Boissy et al. 2016). Par l'empathie, le médecin peut appréhender plus facilement les attentes du patient pour proposer une prise en charge d'un commun accord avec le patient (Stewart et al. 2000). Les compétences en communication peuvent également inciter les patients à participer à leurs soins et à assumer la partie la plus critique de la plupart des plans de soins, l'observance quotidienne (Zolnierek and Dimatteo 2009; Chichirez and Purcărea 2018).

Type d'association		Principale référence	
Processus de soin	Alliance thérapeutique	Chichirez and Purcărea 2018	
	Taux de re-consultation	Cals et al. 2009	
	Attentes du patient	Simpson et al. 2011	
	Relation centrée sur le patient	Stewart et al. 2000	
	Réclamations auprès des médecins	Levinson et al. 1997	
	Conflits éthiques	Shelton et al. 2016	
Résultat pour le patient	Satisfaction	Boissy et al. 2016	
	Résultats cliniques	Qualité de vie	Aiarzaguena et al. 2007
		Douleur	Chassany et al. 2006
		Anxiété et dépression	Girgis et al. 2009
		Perte de poids	Christian et al. 2008
		Pression artérielle	Cooper et al. 2011
		Ionogramme	Sequist et al. 2010

Tableau 2. Association des compétences relationnelles avec la qualité des soins

Les compétences relationnelles sont logiquement associées positivement à de meilleurs indicateurs de santé (Stewart 1995; Kelley et al. 2014) (Tableau 2). Ces compétences ont par exemple montré leur efficacité sur l'anxiété des patients avec comme résultat une diminution

du nombre d'examens d'exploration diagnostique réalisés (Stewart et al. 2000). Les compétences relationnelles sont enfin associées à une diminution des réclamations auprès des établissements de santé (Levinson et al. 1997) et une diminution des conflits éthiques liés aux soins (Shelton et al. 2016).

1.1.3. Importance des compétences relationnelles médecin-patient en formation

Du fait d'un rôle central dans le soin, les compétences relationnelles constituent un enjeu important pour la formation des médecins, aussi bien en formation initiale qu'en formation continue. En effet, ces compétences ne sont pas complètement innées et elles peuvent s'acquérir ou se développer par une formation dédiée (Kurtz et al. 2004). Longtemps mis au second plan de la formation, du fait de la prédominance des connaissances scientifiques dans le cursus, puis des compétences techniques, les compétences non-techniques connaissent un regain d'intérêt depuis quelques dizaines d'années (Junod Perron et al. 2015).

Le développement de l'approche par compétences dans les facultés de médecine s'est d'abord réalisé dans les pays anglo-saxons autour des années 2000 (van Zanten et al. 2007), avant d'arriver plus récemment en France. De multiples programmes pédagogiques ont alors émergé, de manière concomitante avec le développement de la simulation en santé (Sanson-Fisher et al. 2018) et des Examens Cliniques Objectifs Structurés (ECOS) qui ont laissés une large place aux simulations non-techniques (Cömert et al. 2016). En France, la réforme du deuxième cycle des études médicales, actuellement en cours, marque le tournant des études de médecine vers l'approche par compétences. Annoncée dès 2017, cette réforme a permis de faire entrer la communication au cœur des 6 compétences transversales nécessaires à tout étudiant en médecine à la fin du deuxième cycle. Ainsi, un clinicien se doit maintenant d'être "communicateur" (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 - Légifrance). En 2021, les contours des futures épreuves d'ECOS, qui marqueront la fin du deuxième cycle des études de médecine, sont tracés (Arrêté du 21 décembre 2021 relatif à l'organisation des épreuves nationales donnant accès au troisième cycle des études de médecine). Une station d'interrogatoire d'un patient simulé fera partie de cet examen national avec une mise en avant des compétences relationnelles dans les situations cliniques de départ, dont plusieurs situations évoquent une explication, une

information ou une annonce à faire au patient. Dans ce cadre, il est donc indispensable de pouvoir fournir des entraînements et des formations dédiées aux compétences relationnelles au cours des premiers cycles des études médicales. Par ailleurs, le travail des compétences relationnelles participe de fait à l'augmentation de la confiance en soi des étudiants sur le long terme et participe à la diminution des risques psycho-sociaux (Buckley et al. 2016; Omura et al. 2016).

Cette nécessité est d'autant plus importante qu'il a été démontré que l'empathie des étudiants en médecine diminuait au fur et à mesure de leur cursus et de l'exercice médical (Neumann et al. 2011; Leune et al. 2021).

1.2. Rationnel

1.2.1. Évaluation

Considérant la nécessité du développement des compétences relationnelles, un premier défi est de pouvoir évaluer ces compétences. L'évaluation des compétences non-techniques paraît toujours difficile et semble moins fiable qu'une évaluation de compétences techniques. Pourtant, les propriétés psychométriques d'un instrument de mesure d'un concept subjectif peuvent démontrer une meilleure fiabilité qu'une mesure généralement considérée comme objective (Falissard 2008).

Malgré tout, l'évaluation des compétences relationnelles médecin-patient est une tâche complexe. Pour preuve, de nombreuses échelles d'évaluation ont été développées, sans qu'aucune d'entre-elles ne s'impose, avec des contextes d'utilisation différents et des recouvrements incomplets (Boon and Stewart 1998; Schirmer et al. 2005; Zill et al. 2014). Peu d'échelles ont par ailleurs été traduites ou développées spécifiquement en français.

Une première difficulté est la nécessaire prise en compte de la communication non-verbale du médecin (Dyche and Swiderski 2005; Stepanikova et al. 2012). Une deuxième difficulté, dans l'évaluation des compétences relationnelles, est la prise en compte des autres dimensions que la communication, notamment l'empathie ou la relation centrée sur le patient (Di Blasi et al. 2001; Mauksch et al. 2008). Une troisième difficulté est la nécessité d'une échelle d'évaluation qui puisse s'appliquer à de nombreuses situations cliniques, non limitée à une spécialité ou une pathologie, comme c'est fréquemment le cas pour l'oncologie par exemple (Fagerlind et al. 2011; Stubenrauch et al. 2012). Enfin, une dernière difficulté est la prise en compte des caractéristiques de la consultation elle-même, ainsi que de l'évaluateur et du moyen choisi pour

permettre l'évaluation. En effet, plusieurs échelles d'évaluation utilisent l'auto-évaluation (Boon and Stewart 1998) ou la satisfaction du patient (Morel et al. 2018). Toutefois, dans un cadre de formation, une échelle d'hétéro-évaluation par un observateur semble préférable, car bien que l'observation directe des compétences relationnelles par le patient ou le patient standardisé soit une méthode valide, elle ne permet pas d'évaluer l'ensemble des dimensions des compétences relationnelles (Coleman 2000). Ainsi, l'évaluation sur la base d'un enregistrement vidéo de la consultation médicale semble être la méthodologie qui répond le mieux aux inconvénients de l'observation directe (Asan and Montague 2014), en permettant de prendre le temps nécessaire pour l'évaluation et en donnant la possibilité de faire des évaluations avec des observateurs multiples (Ram et al. 1999).

1.2.2. Facteurs associés avec la mesure produite

Il est nécessaire d'identifier et de comprendre les facteurs associés à la mesure produite par l'instrument d'évaluation sélectionné pour adapter les stratégies de développement de ces compétences par la suite. Parmi ceux-ci, certains déterminants peuvent être modifiables et d'autres non (Tableau 3).

Le contexte de la consultation médicale exerce généralement une influence sur les compétences relationnelles. Ainsi, les résultats seront différents en milieu universitaire avec un patient standardisé, qu'en milieu hospitalier avec un patient réel. Les résultats en médecine ambulatoire, notamment en soins primaires, seront sans doute différents également, mais rares sont les études s'étant intéressées à ce contexte particulier de soin (Zill et al. 2014). De même, les sages-femmes sont des professions médicales pour lesquelles les compétences relationnelles spécifiques ne sont que peu étudiées (Aktas and Pasinlioglu 2021) ou de manière concomitante avec les professions infirmières (Alhassan 2019).

D'autres caractéristiques de la consultation ont sans doute une influence sur les compétences relationnelles comme la durée de la consultation et le motif de consultation, sans que ces facteurs aient été spécifiquement démontrés comme corrélés dans ce cadre (Wilson and Childs 2006; Lemon and Smith 2014). L'étude de ces facteurs fait émerger le risque de biais protopathique.

De plus, les caractéristiques du médecin lui-même peuvent influencer la qualité de la relation médecin-patient. Ainsi, faire preuve de responsabilité, de fiabilité ou de comportement éthique

favorisent l'établissement d'une relation de confiance (Mann et al. 2005). Faire preuve d'intelligence émotionnelle permet des soins plus compatissants et empathiques (Lopes et al. 2004; Brackett et al. 2006). Faire preuve de sympathie et sourire au patient facilitent enfin l'échange avec le patient (Hall and Roter 2002).

Chez les étudiants en médecine, de rares études ont établi un lien entre les résultats aux examens universitaires et les scores obtenus à des stations d'interrogatoire dans un contexte de simulation (Murden et al. 2004; Adam et al. 2012).

L'expérience du médecin est également un facteur à prendre en compte (Ruiz Moral et al. 2002). Elle aurait une influence notamment sur la durée de consultation et par extension sur les compétences relationnelles (Wouda and van de Wiel 2012).

Typologies des déterminants	Déterminants	Caractère modifiable	Références
Systémique	Lieu d'exercice : Hôpital vs ambulatoire	Non	Clayton et al. 2011
	Professionnalisation : Formation initiale vs continue	Non	Makoul et al. 2001
Organisationnel	Type de profession médicale	Non	Alhassan 2019
	Type de patients : réels ou standardisés	Oui	van Zanten et al. 2007
	Durée de consultation	Oui	Wouda and van de Wiel 2012
Individuel	Age	Non	Ruiz Moral et al. 2002
	Sexe	Non	Hall and Roter 2002
	Formation aux compétences relationnelles	Oui	Rao et al. 2007
	Résultats aux examens	Oui	Adam et al. 2012
	Caractéristiques psychologiques/sociales	Non	Brackett et al. 2006
	Expérience extra-hospitalière	Oui	

Tableau 3. Typologie des déterminants des compétences relationnelles médecin-patient

Enfin, certaines caractéristiques non modifiables des médecins comme le genre (Tableau 3) ont largement été étudiées. Il a ainsi été montré que les médecins de sexe féminin posaient plus de questions, faisaient plus de retours a posteriori au patient (Hall and Roter 2002). Quant aux patients, ils parlent plus et révèlent plus d'information psychosociale aux médecins de sexe

féminin (Jefferson et al. 2013), expliquant peut-être les meilleures compétences relationnelles relevées généralement chez celles-ci.

1.2.3. Développement des compétences relationnelles

Devant la nécessité de former les professionnels médicaux à ces compétences essentielles, de nombreuses institutions ont mis en place des programmes pédagogiques visant à développer des compétences relationnelles médecin-patient. Bien entendu, les facultés de médecine ont proposé des interventions pédagogiques, principalement centrées sur les compétences en communication (Gilligan et al. 2021), parfois sur l'empathie (Batt-Rawden et al. 2013). Des centres hospitaliers ont mis en place des programmes d'entraînement avec une évaluation en routine des compétences relationnelles (Stein et al. 2005; Gulbrandsen et al. 2013). Le développement des compétences relationnelles en médecine générale reste rare (Fassaert et al. 2007; Burt et al. 2014). A contrario, de nombreux programmes pédagogiques se développent dans des contextes spécifiques comme en chirurgie (Fleetwood et al. 2018), en oncologie (Stubenrauch et al. 2012) ou en dermatologie par exemple (Wang et al. 2015).

De nombreuses études ont étudié l'efficacité d'interventions pédagogiques dédiées au travail de ces compétences, avec des formats et des durées hétérogènes (Rao et al. 2007). La meilleure stratégie pédagogique pour l'amélioration des compétences relationnelles reste cependant incertaine (Junod Perron et al. 2015).

Par ailleurs, l'impact de ces interventions sur l'efficacité réelle de ces dernières restent limitées, car la majorité des études sont descriptives ou sans groupe contrôle (Hulsman et al. 1999). Moins de 2% des articles font référence à des études contrôlées randomisées (Sanson-Fisher et al. 2018). Enfin, les résultats sur le patient et sur la qualité des soins sont rarement décrits (Stein et al. 2005; Zolnieriek and Dimatteo 2009).

1.3. Objectifs

L'objectif général de ce travail était d'actualiser les connaissances sur les outils génériques d'évaluation, sur les déterminants et l'efficacité d'interventions centrées sur les compétences relationnelles médecin-patient, dans le contexte de consultations enregistrées.

Ainsi, les objectifs spécifiques étaient premièrement d'identifier, parmi l'ensemble des échelles d'évaluation existantes, une échelle d'évaluation des compétences relationnelles. L'échelle devait couvrir l'ensemble du champ des compétences relationnelles et s'appliquer à un contexte d'hétéro-évaluation sur la base de consultation médicale vidéo-enregistrée. Elle devait ensuite être valide et fiable pour une utilisation en routine dans les institutions françaises, à la fois hospitalières et universitaires.

Deuxièmement, nous avons tenté d'identifier les facteurs influençant les compétences relationnelles, comme le contexte de soin, les caractéristiques de la consultation et les caractéristiques des médecins.

Enfin, un troisième objectif était d'évaluer l'efficacité des programmes pédagogiques dans le développement des compétences relationnelles, aussi bien en formation initiale, qu'en formation continue. Pour cela, nous avons utilisé des études quasi-expérimentale et expérimentale (essai randomisé).

Pour répondre à ces trois objectifs, nous avons proposé le plan suivant en trois chapitres:

- **Chapitre 1** : Évaluation des compétences relationnelles médecin-patient
 - Revue systématique de la littérature des échelles d'évaluation des compétences relationnelles en consultation médicale
 - Adaptation trans-culturelle et validation des propriétés psychométriques de l'échelle 4-Habits Coding Scheme (4-HCS)
- **Chapitre 2** : Déterminants des compétences relationnelles en consultation médicale
 - Compétences relationnelles en médecine générale ambulatoire
 - Compétences relationnelles en maïeutique
 - Facteurs associés aux compétences relationnelles chez les étudiants en médecins en consultations simulées
 - Impact de la durée de consultation sur les compétences relationnelles
- **Chapitre 3** : Développement des compétences relationnelles par des programmes pédagogiques dédiés
 - Evaluation des dispositifs d'amélioration des compétences relationnelles lors de la formation initiale
 - Protocole d'un essai contrôlé randomisé pour l'évaluation d'un programme pédagogique dédié aux compétences relationnelles en formation continue (EPECREM)
 - Résultats préliminaires de l'étude EPECREM

CHAPITRE 1

Évaluation des compétences relationnelles médecin-patient

2. CHAPITRE 1 : Évaluation des compétences relationnelles médecin-patient

2.1. Partie 1) Revue systématique de la littérature des échelles d'évaluation des compétences relationnelles en consultation médicale

Article soumis — Content and psychometrics properties of standardized instruments measuring physician interpersonal skills in consultation: a Systematic Review

Dans cet article soumis à la revue "Patient Education and Counseling", nous avons réalisé une revue systématique de la littérature visant à identifier des échelles d'évaluation des compétences relationnelles.

Objectif

Nous avons cherché à identifier des instruments standardisés pour évaluer les compétences relationnelles des médecins à partir de consultations vidéo-enregistrées sur vidéo, puis nous avons examiné les données publiées sur les propriétés psychométriques de ces instruments de mesure.

Méthodes

Nous avons interrogé les bases de données électroniques à la recherche d'études publiées entre janvier 1950 et avril 2020. Les études éligibles devaient rendre compte du développement ou de la validation d'instruments standardisés dédiés à l'évaluation des compétences relationnelles des médecins à partir de consultations vidéo-enregistrées de patients adultes. Chaque étude a été évaluée selon les critères du COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) pour déterminer sa qualité.

Résultats

Sur les 6 318 études répertoriées, treize études impliquant neuf instruments standardisés ont été incluses. Les nombres médians de médecins et de participants étaient respectivement de 23 (étendue : 1 à 200) et 71 (étendue : 1 à 950). Sept des neuf instruments étaient multidimensionnels, avec un nombre médian de dimensions de 5 et un nombre médian d'items de 23 (étendue : 7 à 95). Le nombre médian de critères COSMIN remplis dans toutes les études était de 2 sur 10. Le niveau de preuve en termes de validité de contenu, de cohérence interne,

de validité de construit et de fiabilité était limité pour la plupart des instruments standardisés inclus.

Conclusion

Les données étayant les propriétés psychométriques de la plupart des instruments standardisés publiés, destinés à évaluer les compétences relationnelles des médecins lors de consultations médicales vidéo-enregistrées étaient limitées.

Implications pratiques

Nous préconisons l'utilisation de la version américaine originale du 4-Habits Coding Scheme (4-HCS) et de ses adaptations transculturelles, bien que nous recommandions une validation plus poussée de ses propriétés psychométriques.

Validity and reliability of standardized instruments measuring physician communication and interpersonal skills in video-recorded medical consultations – A systematic review

Alexandre Bellier^{1,2,3}, Philippe Chaffanjon³, Joey Fournier³, Noémie Kaladzé³, Alexis Dechosal³, Patrice François^{1,2}, José Labarère^{1,2,3}

¹ Univ. Grenoble Alpes, Computational and Mathematical Biology Team, TIMC, UMR 5525, CNRS, Grenoble, France

² Clinical epidemiology unit, Grenoble Alpes University Hospital, Grenoble, France

³ Medical School, Univ. Grenoble Alpes, Grenoble, France

Corresponding author: Alexandre Bellier, Clinical epidemiology unit, Grenoble Alpes University Hospital, CS10217, 38043 Grenoble Cedex 09, France. Tel.: +33 (0)4 76 76 87 67.
E-mail address: abellier@chu-grenoble.fr

Highlights

- Various instruments measuring physician interpersonal skills have been published.
- Nine standardized instruments designed to assess interpersonal skills throughout video-recorded medical consultations were identified.
- Evidence on psychometric attributes for these instruments was limited due to inappropriate primary study design or reporting.
- We advocate the use of the 4-Habits Coding Scheme (4-HCS) and its cross-cultural adaptations to assess physician interpersonal skills during video-record medical consultations.

ABSTRACT

Objective

To identify standardized instruments measuring interpersonal physician skills based on video-recorded consultation and to review published evidence on their psychometric properties.

Methods

We searched electronic databases for primary studies published from January 1950 to April 2020. Eligible studies had to report on the development or validation of standardized instruments dedicated to the assessment of physician interpersonal skills based on video-recorded consultations with adult patients. Each study was rated according to COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) criteria.

Results

Of 6,318 studies retrieved, 13 primary studies involving nine standardized instruments were included. The median numbers of physicians and participants were 23 (range, 1 to 200) and 71 (range, 1 to 950), respectively. Seven out of nine instruments were multidimensional, with a median number of dimensions of 5 and a median number of items of 23 (range, 7 to 95). The median number of COSMIN criteria fulfilled across studies was 2 out of 10. The level of evidence on content validity, internal consistency, construct validity and reliability was limited for most instruments.

Conclusion

Limited evidence supports the psychometric attributes for most published standardized instrument dedicated to assess physician interpersonal skills throughout video-recorded medical consultations.

Practice implications

We advocate the use of the original US version of the 4-Habits Coding Scheme (4-HCS) and its cross-cultural adaptations and recommend further validation of its psychometric attributes.

Keywords Interpersonal skills, Medical consultation, Psychometric properties, Review, Standardized instrument

1. Introduction

Physician communication and interpersonal skills encompass the ability to collect clinically relevant information, understand patient concerns, provide appropriate counselling and therapeutic instructions, and establish caring relationships with patients [1–3]. They are key components of effective medical consultations [4] and represent physician qualities that are most desired by patients [5]. Accumulating evidence exists that high-quality communication contributes to enhanced patient satisfaction [6], greater adherence to treatment [7], better health outcomes [8,9], and decreased risk of malpractice claims [10]. Many organizations have therefore implemented structured training programs and routinely assessed clinician communication skills [6,11,12].

Assessing physician-patient interaction during medical consultations is a complex task [13], because communication is not limited to the exchange of verbal information but also includes non-verbal expression (*e.g.*, tone, facial expression, eye contact) [14–16]. Additionally, less tangible aspects of communication, often referred as interpersonal skills, may convey physician expression of humanistic attitudes like understanding, supportiveness, and empathy [5,14,17,18].

Although direct evaluations of physician communication skills by standardized patients or peers are valid assessment methods, they may not capture all components of communication in real time [19] and requires considerable resource investments [20].

Video-recording of consultations with real or standardized patients has the potential for overcoming some of the challenges of direct observation [21]. Video-recording captures all events that occurred during consultations and allows raters to review consultations as many times as necessary [21]. Video-recording can also be rated by several observers without consultations being disrupted [20]. Additionally, video-recording allows for providing physicians with feedback on their own performance [5,21,22].

Four reviews of standardized instruments intended to assess clinician communication and interpersonal skills were published between 1977 and 2014 [13,23–25]. Yet, these reviews have not been updated, did not use formal literature search strategies [13], included instruments with various data acquisition methods (*i.e.*, patient-reported measures, real-time observation, audio-recording, or video-recording) [23,24], or were restricted to rating scales for assessing medical student communication skills as part of objective structured clinical examinations [25]. To our knowledge, a

systematic review focusing on instruments dedicated to assessment of communication and interpersonal skills based on video-recorded medical consultations is still lacking.

In this context, we conducted a systematic review of published standardized instruments or questionnaires designed to assess physician communication and interpersonal skills in video-recorded medical consultations. More specifically, we investigated the underlying conceptual framework, we appraised the development and validation processes, we examined the validity and reliability estimates, and we assessed the transparency of reporting and the level of evidence supporting the psychometric attributes for each instrument, according to current standards.

2. Methods

This systematic review was performed according to current guidelines [26,27] and complied with the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA) statement [28]. The rationale and methods were prespecified and reported in a protocol that was registered with PROSPERO (CRD42017060926).

2.1. Eligibility criteria

Study design. Primary studies reporting on the development or validation of standardized instruments intended to measure physician communication or interpersonal skills were eligible in this systematic review. These included prospective or cross-sectional studies with random or non-random (*e.g.*, consecutive) samples of participants. Qualitative studies were not in the scope of this review.

Data acquisition. Our review was restricted to primary studies using video-recorded data collection methods, although we anticipated various methods of data acquisition [23]. More specifically, primary studies using patient- or self-reported measures, real-time observation, and audio-recording were not eligible for this review.

Setting. We focused on instruments designed to be used in the primary care consultation setting with adult patients (≥ 18 years). Studies conducted in the medical education setting with real or standardized patients were also considered for inclusion. Consistent with Zill et al. [24], we included generic instruments applicable to a broad range of health conditions and patients.

Conversely, condition-, age-, or gender-specific instruments were excluded. Instruments targeting specific dimensions

of communication or interpersonal skills were also excluded.

Participants. Primary studies reporting on communication or interpersonal skills for senior physicians or medical students were eligible. Conversely, instruments measuring communication skills for other healthcare providers (e.g., nurses, pharmacists) or instruments measuring communication within the healthcare provider team were excluded.

Applied inclusion and exclusion criteria are displayed in Table 1.

Inclusion criteria	
1	The aim of study is to assess a competence measure instrument
2	The measured construct corresponds to interpersonal skills (relational and communication competences)
3	The target group is adult patients
4	The relationship is between patient and physician (including medical students)
5	The relationship is evaluated based on video-recorded consultations
Exclusion criteria	
1	Patients are < 18 years
2	The instrument is condition-specific or restricted to a subgroup of patients
3	Measure is self-reported
4	Document type is a systematic review

Table 1. Inclusion and exclusion criteria

2.2. Search strategy

Electronic databases. Studies were identified by searching Medline via PubMed, EMBASE via Ovid, PsycINFO, and Web of Science. The last search was conducted on April 08, 2020.

Other information sources. We supplemented the electronic search by scanning the reference lists of retrieved original articles and previously published review articles for additional studies.

Strategy. Electronic search strategies were developed by a member from the project team (AB) and critically reviewed by a health sciences librarian, using the Peer Review of Electronic Search Strategies (PRESS)-2015 Guideline Statement [29]. The search concepts included 1) physician-patient communication or interpersonal skills, 2) consultation, and 3) rating scale or instrument. We used both standardized medical subject heading (MeSH) and text words (Appendix 1). No type of document restriction was applied and no methodology filters were used. Adjustments were made to account for differences in syntax and subject headings (EMTREE) across electronic bibliographic databases. A member from the project team with experience of systematic reviews executed electronic searches. Additional searches were performed in each database, using the names of the instruments which were identified at the initial search.

Language. Electronic literature searches were conducted without language restrictions. However full text articles published in languages other than English or French were not assessed for eligibility because of limited resources for translation. For transparency purpose, a list of potentially relevant articles published in other languages was provided in Appendix 2.

2.3. Study records

Data management. Literature search results were uploaded along with titles and abstracts into a reference management software (Zotero®). We identified duplicate publications reporting data from the same study by comparing author names, study sites, and sample sizes.

Selection process. One review author (AB) screened the titles and abstracts, where available, of all records yielded by the literature search against pre-specified eligibility criteria. The reviewer assessed whether the citation 1) reported data from an original research study, 2) focused on the development or validation of a standardized instrument or questionnaire 3) evaluated physician communication or interpersonal skills in medical consultations, and 4) used video-recorded data collection methods. Full-text articles of records identified as

being potentially relevant were assessed. Reasons for excluding study records were collected. The reviewer was not blinded to authors, date and journal of publication.

2.4. Data extraction

A review author (AB) performed qualitative and quantitative data extraction. For this purpose, a data extraction form was developed, pretested, and adjusted as necessary. Details of study (first author name, publication year), instrument (name, version, number of items, number of subscales, response scale, language, validity and reliability estimates), and study population (age, gender, condition, setting, country, enrollment period, sample size, recruitment) characteristics were extracted from published articles.

2.5. Assessment of study quality

Two authors (AB and JF) performed independent quality assessment of primary studies, using a tool adapted from the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist [27,30,31]. The COSMIN checklist comprises ten boxes addressing methodological standards for reliability and validity properties of measurement instruments. Each box includes 4 to 35 items rated on a four-point scale (i.e.,

inadequate, doubtful, adequate, or very good) [27,30,31].

Eight measurement properties of instruments were appraised, including: (1) internal consistency, which reflected the uni- or multi-dimensionality of the instrument and the interrelatedness among its items; (2) reliability, which corresponded to the proportion of total measurement variance explained by true variations across patients; (3) measurement error, reflecting the systematic and random error in an instrument measurement that was not attributable to true changes in the measured construct; (4) content validity, which designated the degree to which the content of an instrument was an adequate reflection of the construct to be measured; (5) structural validity, designating the degree to which the observed empirical dimensionality was an adequate reflection of the hypothesized instrument dimensionality; (6) construct (criterion) validity, designating the relationships between an instrument measure and other measures, in accordance with prespecified hypotheses; (7) responsiveness, designating the ability of an instrument to detect change over time in the construct to be measured; (8) cross-cultural validity, indicating the degree to which the performance of a cross-cultural adaptation reflected the performance of the original instrument [27,30,31].

Consistent with Terwee et al. [32], we evaluated internal consistency using Cronbach's alpha. We assessed inter-rater and test-retest reliability using intra-class correlation coefficient for continuous variables or weighted Cohen's Kappa coefficient for ordinal variables, respectively. Pearson or Spearman correlation coefficients were reported if used, although they were not adequate measures of reliability [32]. We considered a study fulfilled the criterion for internal consistency when factor analysis was applied and Cronbach's alpha was ≥ 0.70 [33]. Reliability was deemed satisfactory when the intra-class correlation coefficient was ≥ 0.70 [32,33].

Cross-cultural adaptations were appraised according to Guillemin et al. [34]. These guidelines advocate back-translation techniques and committee review, pre-testing techniques and re-examining the weights of scores, in order to ensure semantic and conceptual equivalence [34].

2.6. Data analysis

Information on key primary study, participant, and standardized instrument characteristics were summarized in tables. Assessment of methodological quality across primary studies was tabulated. Descriptive summary statistics were used for reporting continuous (mean and

standard deviation or median and range, where appropriate) and categorical (numbers and percentages) variables. Statistical analysis was performed using Stata Special Edition version 16.0 (Stata Corporation, College Station, TX, USA).

3. Results

3.1. Literature search

A total of 6,318 unique citations were identified through database searching (Figure 1). Twenty-five additional records were identified from citation and reference tracking. After removing 1,749 identical records identified from different sources and 10 duplicates, the title and abstract for 4,584 records were screened for eligibility. Of these, 108 records were identified as being potentially relevant and full-text article was retrieved for more thorough eligibility review. After excluding 95 citations that did not fulfill the inclusion criteria based on the full-text article, our systematic review included 13 primary studies involving nine standardized instruments.

3.2. Primary study characteristics

Eight studies were conducted in Western Europe [35–42], three in the US and two in Canada [2,43–46]. Five studies were carried out in general practitioner healthcare facilities [37,38,40,41,45], while the other studies were conducted in hospital

outpatient clinics or at university facilities. Ten studies were performed with real patients and three studies involved standardized patients.

The median numbers of physicians and participants were 23 (range, 1 to 200) and 71 (range, 1 to 950), respectively. The median percentage of male physicians was 54.5% (range, 28.6% to 75.5%) and the median percentage of male patients was 42.9% (range, 37.3% to 50.0%), across nine studies for which the information was available (Table 2).

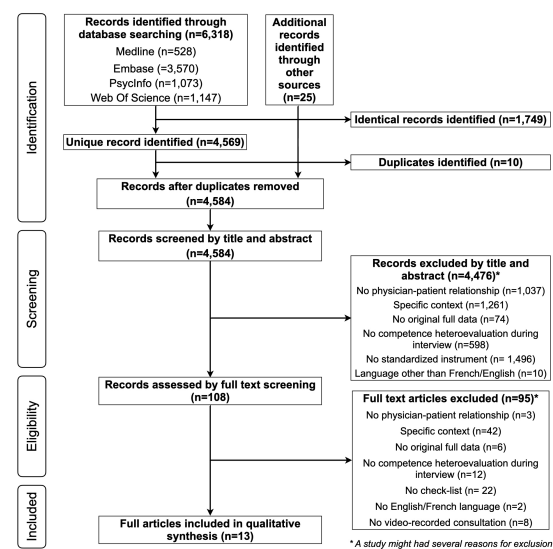


Fig 1. PRISMA flowchart

3.3. Standardized instrument characteristics

All but two standardized instruments were developed in English language. The exceptions were a Dutch language questionnaire [39] and a Canadian French language questionnaire [45]. Cross-cultural adaptations of the 4-Habits Coding Scheme

(4-HCS) in French, German and Norwegian were validated using video-recorded consultations (Table 3). The conceptual framework was reported for two standardized instruments [2,41] and was lacking or unspecified for all other instruments. Seven instruments were multidimensional, with a median number of five dimensions (range, 2 to 6) while this information was not available for two instruments. Item generation used literature reviews (n = 2) [2,45], expert opinions (n = 2) [45,46], focus groups (n = 2) [2,41], or key informant interviews (n = 2) [2,41], with two or more approaches combined for three instruments. The median number of items was 23 (range, 7 to 95), with substantial overlap in item content across identified instruments (Appendix 3).

3.4. Psychometric attributes of standardized instruments

Internal consistency as assessed by Cronbach's alpha ranged from 0.50 to 0.96 across eight studies (Table 4). Only two studies reported on inter-item correlations [35,42] and four studies reported on factorial analysis [2,35,39,45]. Exploration of content validity or construct validity was missing in 3 and 9 of the 13 studies, respectively. Inter-rater reliability was adequately quantified by intra-class correlation coefficient for five studies

(range, 0.36 to 0.96). Other studies measured inter-rater reliability by using

Cohen's Kappa coefficient (n = 2) or Pearson's correlation coefficient (n = 3).

Only two studies reported on intra-rater reliability, using intra-class correlation coefficient [35,42].

3.5. Methodological quality of primary studies

No primary study reported information on development (Box 1), measurement error (Box 7), and responsiveness (Box 10) (Table 5). Only one study [35] complied with each of the seven remaining COSMIN criteria. None of the studies was based on Item Response Theory. Internal consistency (Box 4) was reported for eight studies [2,35,38,39,42–45]. Inter-rater reliability estimates (Box 6) were reported adequately for nine studies [35–40,42–44,46]. Content validity (Box 2) was reported for four studies [2,41,45,46] with only two studies [2,41] assigned a very good rating. Criterion validity (Box 8) was available for four studies [35,39,42,43], with three of them rated as very good [35,39,43]. Adequate or very good level of evidence on cross-cultural validity was available for two out of four studies reporting information on this item (Box 5) [35,42].

Measure	Authors (Year)	Study characteristics					Participant characteristics						
		Country	Number of centers	Setting	Type of consultations	Type of physicians	Type of patients	Number of physicians	Male physician percentage	Number of patients	Patient mean age	Male patient percentage	Number of consultations
4-HCS	Bellier (2020)	France	1	University	OSCE	Students	Standardized patients	200	...	20	200
	Scholl (2014)	Germany	1	Outpatient clinics	Routine medical visits	General Practitioner, Specialists	Real patients	22	54.5	67	55.7	37.3	69
	Fossli Jensen (2010)	Norway	...	Outpatient clinics, hospital	Routine medical visits	Specialists	Real patients	71	62.0	497	46	48.0	497
	Krupat (2006)	USA	1	Outpatient clinics, hospital	Routine medical visits	Specialists and Residents	Real patients	50	50.0	100	62	50.0	100
GCRS	Burt (2014)	United Kingdom	...	GP healthcare facility	Routine medical visits	General Practitioner	Standardized patients	23	...	42	42
ALOS-global	Fassaert (2007)	Netherlands	...	GP healthcare facility	Routine medical visits	General Practitioner	Real patients	139	75.5	524	46.09	42.9	524
PCBI	Zandbelt (2005)	Netherlands	...	GP healthcare facility	Routine medical visits	General Practitioner, Specialists, Residents	Real patients	30	53.0	330	53	42.0	323
LIV-MAAS	Enzer (2003)	United Kingdom	...	GP healthcare facility	Routine medical visits	General Practitioner	Real patients	8	50.0	71	...	44.0	71
	Robinson (2002)	United Kingdom	1	GP healthcare facility	Routine medical visits	General Practitioner	Real patients	15	66.7	213	39.5	38.0	213
RCS-O	Gallagher (2001)	USA	1	University	OSCE	Students	Standardized Patients	13	61.5	20	20
CAT	Makoul (2001)	USA	1	Outpatient clinics, hospital	Routine medical visits	General Practitioner, Specialists, Residents and Students	Real patients	38	...	950	950
ECRMP	Côté (2001)	Canada	29	Ambulatory and Hospital	Routine medical visits	General Practitioner	Real patients	1	...	1	1
Verdun Checklist	Lehmann (1990)	Canada	11	...	Routine medical visits	Residents	Real patients	14	28.6	28	28

Table 2. Primary study and participant characteristics

Abbreviations: 4-HCS = 4-Habits Coding Scheme; ALOS-global = Active Listening Observation Scale; CAT = Communication Assessment Tool; ECRMP = "Evaluation des compétences relationnelles du médecin avec le patient"; GCRS = Global Consultation Rating Scale; LIV-MAAS = Liverpool-MAAS; NA = Not Applicable; OSCE = Objective Structured Clinical Examination; PCBI = Patient-centred Behaviour Coding Instrument; RCS-O = Relational Communication Scale for Observational Measurement

Measure	Authors (Year)	Language	Number of dimensions	Item generation	Number of items	Response format
4-HCS	Bellier (2020)	French		NA		
	Scholl (2014)	German		NA		
	Fossli Jensen (2010)	English	4	NA	23	5-point Likert Scale
	Krupat (2006)	English		...		
GCRS	Burt (2014)	English	12	...	37	3-point Scale
ALOS-global	Fassaert (2007)	English	Unspecified	...	7	5-point Likert Scale
PCBI	Zandbelt (2005)	Dutch	2	...	19	Other
LIV-MAAS	Enzer (2003)	English		NA		
	Robinson (2002)	English	6	Focus group and key informant interviews	95	3-Point Scale and Nominal Scale
RCS-O	Gallagher (2001)	English	6	...	34	5-point Likert Scale
CAT	Makoul (2001)	English	5	Literature review, focus group and key informant interviews	15	5-point Likert Scale
ECRMP	Côté (2001)	French Canadian	2	Literature review and expert opinions	15	4-Point Likert Scale
Verdun Check-list	Lehmann (1990)	English and French Canadian	Unspecified	Expert opinions	25	Nominal scale (dichotomous)

Table 3. Characteristics of standardized instruments

Abbreviations: 4-HCS = 4-Habits Coding Scheme; ALOS-global = Active Listening Observation Scale; CAT = Communication Assessment Tool; ECRMP = “Evaluation des compétences relationnelles du médecin avec le patient”; GCRS = Global Consultation Rating Scale; LIV-MAAS = Liverpool-MAAS; PCBI = Patient-centred Behaviour Coding Instrument; RCS-O = Relational Communication Scale for Observational Measurement

Measure	Authors (Year)	Internal consistency		Inter-rater reliability	Intra-rater reliability	Content validity	Construct validity	Factor analysis
		Cronbach's alpha	Inter-items correlation					
4-HCS	Bellier (2020)	0.72-0.88	0.29-0.90	ICC: 0.74	ICC: 0.85	YES	YES	YES
	Scholl (2014)	0.38-0.51	0.01-0.52	ICC: 0.36-0.96	ICC: 0.00-1.00	...	YES	...
	Fossli Jensen (2010)	ICC: 0.71-0.87
	Krupat (2006)	0.51-0.81	...	Correlation coefficient: 0.69-0.80	YES	...
GCRS	Burt (2014)	Correlation coefficient: 0.88
ALOS-global	Fassaert (2007)	0.84	...	Kappa: 0.52	...	YES	YES	...
PCBI	Zandbelt (2005)	0.50-0.64	...	ICC: 0.53-0.93	...	YES	YES	YES
LIV-MAAS	Enzer (2003)	ICC: 0.69-0.91
	Robinson (2002)	YES
RCS-O	Gallagher (2001)	0.60-0.95	...	Kappa: 0.02-0.72
CAT	Makoul (2001)	0.96	YES	...	YES
ECRMP	Côté (2001)	0.92	YES	...	YES
Verdun Check-list	Lehmann (1990)	Correlation coefficient: 0.39	...	YES

Table 4. Psychometric properties of standardized instruments

Abbreviations: 4-HCS = 4-Habits Coding Scheme; ALOS-global = Active Listening Observation Scale; CAT = Communication Assessment Tool; ECRMP = "Evaluation des compétences relationnelles du médecin avec le patient"; GCRS = Global Consultation Rating Scale; LIV-MAAS = Liverpool-MAAS; PCBI = Patient-centred Behaviour Coding Instrument; RCS-O = Relational Communication Scale for Observational Measurement

4. Discussion and Conclusion

4.1. Discussion

This comprehensive systematic review identified nine standardized instruments dedicated to measuring physician

communication and interpersonal skills during video-recorded medical consultations. Overall, evidence on the validity and reliability for these instruments relies on 13 primary studies totaling 2,633 medical consultations and 698 physicians in

North America and Western Europe. The identified instruments varied in purposes, content, dimensions, wording, format of response, and presentation. Most studies did not comply with current standards for transparent reporting of measurement scale development and validation.

4.2. Instruments

Because no universal definition of physician communication and interpersonal skills exists, their measurement is indirect, relying on observable behaviors that are considered relevant to the concept under study by the instrument developers [47]. It is therefore important for the potential user of an instrument to be aware of the underlying conceptual framework [48]. Yet most instruments identified in our systematic review were published with no reference to the concept that they were intended to measure, except in general terms. Hence, the results that they produce may be difficult to interpret without an adequate conceptual foundation [49]. Exceptions to this observation were the CAT, LIV-MASS and 4-HCS. This last standardized instrument relies on the “Four Habits Model” which refers to basic medical interview tasks that are organized within Internal consistency was rated very good for most instruments, based on Cronbach’s alpha coefficient. Inter-rater reliability is

four dimensions, as part of a training program.

The literature suggests that physician communication and interpersonal skills measured during medical consultations is a multidimensional construct [24]. Accordingly, the identified instruments delineated a total of 27 different dimensions of communication and interpersonal skills. Yet, empirical correlations between prespecified dimensions were high and factorial analysis departed from the hypothesized multidimensional structure for several instruments. These findings do not necessarily invalidate the postulated structure of the instruments although further investigating the factorial structure is warranted [50]. The number of dimensions was not reported for three instruments [37,38,46]. Assuming that they were unidimensional, these instruments produce single overall scores, giving equal weight to various aspects of consultation - an assumption that deserve further investigation. Additionally, unidimensional instruments may mask divergent performance on different tasks of medical consultations. For multidimensional instruments, we therefore recommend that subscale scores be computed separately. critical when measuring physician communication and interpersonal skills using video-recorded consultations. Only

seven instruments [35,36,38–40,42,43] demonstrated satisfactory level of reliability quantified by intra-class correlation. We recommend using two independent raters with some instruments, including the 4-HCS, in routine practice [35].

The primary studies identified encompass a broad range of clinicians, consultations, patients (ever real or standardized), and organizations. However, each instrument

Measure	Authors (Year)	Content validity (Box 2)	Structural validity (Box 3)	Internal consistency (Box 4)	Cross-cultural validity (Box 5)	Inter-rater reliability (Box 6)	Criterion validity (Box 8)	Hypotheses testing for construct validity (Box 9)
4-HCS	Bellier et al. (2020)	Doubtful	Very good	Very good	Very good	Very good	Very good	Very good
	Scholl et al. (2012)	...	Doubtful	Very good	Adequate	Very good	Adequate	Very good
	Fossli	...	Inadequate	Inadequate	Inadequate			Inadequate
	Jensen et al. (2010)					Very good	Inadequate	
GCRS	Krupat et al. (2006)	Doubtful	Inadequate	Very good	NA	Very good	Very good	Adequate
	Burt et al. (2014)	Doubtful	Inadequate	Inadequate	NA	Adequate	Inadequate	Inadequate
ALOS-global	Fassaert et al. (2007)	Doubtful	Adequate	Very good	NA	Very good	Inadequate	Adequate
PCBI	Zandbelt et al. (2005)	Doubtful	Inadequate	Very good	NA	Very good	Very good	Doubtful
LIV-MAAS	Enzer et al. (2003)	...	Inadequate	Inadequate	...	Very good	Inadequate	Inadequate
	Robinson et al. (2002)	Very good	Inadequate	Inadequate	NA	Doubtful	Inadequate	Inadequate
RCS-O	Gallagher et al. (2001)	Doubtful	Inadequate	Very good	NA	Adequate	Inadequate	Inadequate
CAT	Makoul (2001)	Very good	Very good	Very good	NA	Doubtful	Inadequate	Inadequate
ECRMP	Côté et al. (2001)	Adequate	Adequate	Very good	NA	Doubtful	Inadequate	Inadequate
Verdun Checklist	Lehmann et al. (1990)	Adequate	Inadequate	Inadequate	NA	Doubtful	Inadequate	Inadequate

Table 5. Level of evidence on psychometric properties for standardized instruments according to COSMIN criteria

The following criteria were undocumented for any standardized instrument: instrument development (Box 1), Measurement error (Box 7) and Responsiveness (Box 10)

Abbreviations: 4-HCS = 4-Habits Coding Scheme; ALOS-global = Active Listening Observation Scale; CAT = Communication Assessment Tool; COSMIN = COnsensus-based Standards for the selection of health status Measurement Instruments; ECRMP = “Evaluation des compétences relationnelles du médecin avec le patient”; GCRS = Global Consultation Rating Scale; LIV-MAAS = Liverpool-MAAS; NA = Not Applicable; PCBI = Patient-centred Behaviour Coding Instrument; RCS-O = Relational Communication Scale for Observational Measurement

was developed in a single country, healthcare system, or setting, an observation that considerably limits the external validity of the primary study findings and the transportability of the standardized instruments. A noticeable exception is the 4-HCS, which was originally developed within the US Kaiser Permanente Health Maintenance

Organization and subsequently cross-culturally adapted in Norway, Germany and France.

4.3. Primary studies

Because of poor to fair methodological quality of primary studies according to COSMIN criteria [27,30,31], most instruments lacked adequate validation, thus hindering the confidence in their attributes. Six studies [37,40,42,44–46] did not have an adequate sample size of 100 participants or more, as advocated by the COSMIN group [31].

Three COSMIN criteria, including development, responsiveness and measurement error, were not available for any instruments identified by our review [31]. No study reported on the entire set of psychometric properties. Only four studies explored both content and construct validity [2,35,39,45] and investigated empirical factor structure, as recommended [33].

Studies that received the highest quality rating involved the original US version 4-HCS [43], the cross-cultural adaptations of 4-HCS in French [35] and German [42], the PCBI [39] and CAT [2], respectively.

Our results are consistent with previous reviews [4,13,23,24] although they were focused on physician-patient communication. In contrast, our systematic review addresses physician interpersonal skills, which is a broader concept that encompasses other dimensions including communication, empathy, and the structuring of medical consultation [12]. Previous reviews did not formally evaluate the psychometric properties of instruments [4] or were restricted to instruments developed for the use in medical education settings [23]. This explains why we identified 4 additional articles that were not included in previous reviews [35,38,44,45]. To our knowledge, our systematic review is the first addressing instruments designed to rate video-recorded medical consultations.

4.4. Limitations

The strengths of our review include the performance of search strategy according to PRESS Guideline Statement [29], the compliance with current standards for systematic reviews [26], and the assessment of the methodological quality of primary studies by two independent raters using

most recent update of COSMIN checklist [27,31].

However, our systematic review also has limitations that must be considered. First, the interpretation of the findings is inevitably limited by substantial between-study heterogeneity. The observed variations in internal consistency and reliability across instruments might reflect true variations in study participant characteristics or settings. Second, the search strategy was limited to studies published in English or French and all identified instruments were developed in North America or Western Europe. Hence our findings may not apply to other regions although potentially relevant articles published in other languages were listed in Appendix 2. Third, literature search was performed by a single author.

4.5. Conclusion

This systematic review highlights the poor methodological quality and the lack of transparency for primary studies reporting the development of standardized instruments designed to measure physician communication and interpersonal skills in video-recorded consultations. The nine identified instruments cannot be used interchangeably, and the potential user should examine carefully the underlying conceptual framework and the

psychometric properties for guiding his choice.

4.6. Practice Implications

This systematic review provides an overview of published standardized instruments measuring physician interpersonal skills. The results from this study can help researchers or managers identify the most appropriate instrument for assessing physician interpersonal skills during video-recorded medical consultation. We advocate the use of the 4-HCS in routine practice although further validation is warranted, depending on the setting.

Author contributions

AB, PC, NK, AD, PF and JL contributed to the design of the review. AB and JF assessed the quality of included studies. AB and JL drafted the manuscript and all authors revised and approved its final version.

Funding

This research did not receive any specific grant from funding agencies in the public commercial, or not-for-profit sectors.

Conflict of interest

The authors performed the cross-cultural adaptation of the 4-HCS in French.

References

- [1] M. van Zanten, J.R. Boulet, D. McKinley, Using standardized patients to assess the interpersonal skills of physicians: six years' experience with a high-stakes certification examination, *Health Commun.* 22 (2007) 195–205. <https://doi.org/10.1080/10410230701626562>.
- [2] G. Makoul, The SEGUE Framework for teaching and assessing communication skills, *Patient Educ. Couns.* 45 (2001) 23–34.
- [3] J.F. Ha, N. Longnecker, Doctor-patient communication: a review, *Ochsner J.* 10 (2010) 38–43.
- [4] L.M. Ong, J.C. de Haes, A.M. Hoos, F.B. Lammes, Doctor-patient communication: a review of the literature, *Soc. Sci. Med.* 1982. 40 (1995) 903–918.
- [5] S.A. Simpson, C. Shaw, R. McNamara, What is the most effective way to maintain weight loss in adults?, *BMJ.* 343 (2011) d8042.
- [6] A. Boissy, A.K. Windover, D. Bokar, M. Karafa, K. Neuendorf, R.M. Frankel, J. Merlino, M.B. Rothberg, Communication Skills Training for Physicians Improves Patient Satisfaction, *J. Gen. Intern. Med.* 31 (2016) 755–761. <https://doi.org/10.1007/s11606-016-3597-2>.
- [7] K.B.H. Zolnieriek, M.R. Dimatteo, Physician communication and patient adherence to treatment: a meta-analysis, *Med. Care.* 47 (2009) 826–834. <https://doi.org/10.1097/MLR.0b013e31819a5acc>.
- [8] J. Stewart, Northern practices offer a challenge but few MDs willing to make the move, *CMAJ Can. Med. Assoc. J. J. Assoc. Medicale Can.* 153 (1995) 667–668.
- [9] J.M. Kelley, G. Kraft-Todd, L. Schapira, J. Kossowsky, H. Riess, The influence of the patient-clinician relationship on healthcare outcomes: a systematic review and meta-analysis of randomized controlled trials, *PloS One.* 9 (2014) e94207. <https://doi.org/10.1371/journal.pone.0094207>.
- [10] W. Levinson, D.L. Roter, J.P. Mullooly, V.T. Dull, R.M. Frankel, Physician-patient communication. The relationship with malpractice claims among primary care physicians and surgeons, *JAMA.* 277 (1997) 553–559.
- [11] R. Brown, S. Dunn, K. Byrnes, R. Morris, P. Heinrich, J. Shaw, Doctors' stress responses and poor communication performance in simulated bad-news consultations, *Acad. Med. J. Assoc. Am. Med. Coll.* 84 (2009) 1595–1602. <https://doi.org/10.1097/ACM.0b013e3181baf537>.

- [12] T. Stein, R.M. Frankel, E. Krupat, Enhancing clinician communication skills in a large healthcare organization: a longitudinal case study, *Patient Educ. Couns.* 58 (2005) 4–12. <https://doi.org/10.1016/j.pec.2005.01.014>.
- [13] J.M. Schirmer, L. Mauksch, F. Lang, M.K. Marvel, K. Zoppi, R.M. Epstein, D. Brock, M. Pryzbylski, Assessing communication competence: a review of current tools, *Fam. Med.* 37 (2005) 184–192.
- [14] L. Dyche, D. Swiderski, The effect of physician solicitation approaches on ability to identify patient concerns, *J. Gen. Intern. Med.* 20 (2005) 267–270. <https://doi.org/10.1111/j.1525-1497.2005.40266.x>.
- [15] L. Dyche, Interpersonal skill in medicine: the essential partner of verbal communication, *J. Gen. Intern. Med.* 22 (2007) 1035–1039. <https://doi.org/10.1007/s11606-007-0153-0>.
- [16] I. Stepanikova, Q. Zhang, D. Wieland, G.P. Eleazer, T. Stewart, Non-verbal communication between primary care physicians and older patients: how does race matter?, *J. Gen. Intern. Med.* 27 (2012) 576–581. <https://doi.org/10.1007/s11606-011-1934-z>.
- [17] L.B. Mauksch, D.C. Dugdale, S. Dodson, R. Epstein, Relationship, communication, and efficiency in the medical encounter: creating a clinical model from a literature review, *Arch. Intern. Med.* 168 (2008) 1387–1395. <https://doi.org/10.1001/archinte.168.13.1387>.
- [18] Z. Di Blasi, E. Harkness, E. Ernst, A. Georgiou, J. Kleijnen, Influence of context effects on health outcomes: a systematic review, *Lancet Lond. Engl.* 357 (2001) 757–762.
- [19] T. Coleman, Using video-recorded consultations for research in primary care: advantages and limitations, *Fam. Pract.* 17 (2000) 422–427.
- [20] P. Ram, R. Grol, J.J. Rethans, B. Schouten, C. van der Vleuten, A. Kester, Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility, *Med. Educ.* 33 (1999) 447–454.
- [21] O. Asan, E. Montague, Using video-based observation research methods in primary care health encounters to evaluate complex interactions, *Inform. Prim. Care.* 21 (2014) 161–170. <https://doi.org/10.14236/jhi.v21i4.72>.
- [22] S. Nilsen, A. Baerheim, Feedback on video recorded consultations in medical teaching: why students loathe and love it - a focus-group based qualitative study, *BMC Med. Educ.* 5 (2005) 28. <https://doi.org/10.1186/1472-6920-5-28>.
- [23] H. Boon, M. Stewart, Patient-physician communication assessment instruments: 1986 to 1996 in review, *Patient Educ. Couns.* 35 (1998) 161–176. [https://doi.org/10.1016/S0738-3991\(98\)00063-9](https://doi.org/10.1016/S0738-3991(98)00063-9).
- [24] J.M. Zill, E. Christalle, E. Müller, M. Härter, J. Dirmaier, I. Scholl, Measurement of

- physician-patient communication--a systematic review, *PloS One*. 9 (2014) e112637. <https://doi.org/10.1371/journal.pone.0112637>.
- [25] M. Cömert, J.M. Zill, E. Christalle, J. Dirmaier, M. Härter, I. Scholl, Assessing Communication Skills of Medical Students in Objective Structured Clinical Examinations (OSCE)--A Systematic Review of Rating Scales, *PloS One*. 11 (2016) e0152717. <https://doi.org/10.1371/journal.pone.0152717>.
- [26] J. Higgins, S. Green, *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.1.0, The Cochrane Collaboration, 2011.
- [27] L.B. Mokkink, C.B. Terwee, D.L. Patrick, J. Alonso, P.W. Stratford, D.L. Knol, L.M. Bouter, H.C.W. de Vet, The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study, *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* 19 (2010) 539–549. <https://doi.org/10.1007/s11136-010-9606-8>.
- [28] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, PRISMA Group, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *Ann. Intern. Med.* 151 (2009) 264–269, W64. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>.
- [29] J. McGowan, M. Sampson, D.M. Salzwedel, E. Cogo, V. Foerster, C. Lefebvre, PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement, *J. Clin. Epidemiol.* 75 (2016) 40–46. <https://doi.org/10.1016/j.jclinepi.2016.01.021>.
- [30] L.B. Mokkink, C.B. Terwee, D.L. Patrick, J. Alonso, P.W. Stratford, D.L. Knol, L.M. Bouter, H.C.W. de Vet, The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes, *J. Clin. Epidemiol.* 63 (2010) 737–745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>.
- [31] L.B. Mokkink, C.B. Terwee, D.L. Knol, P.W. Stratford, J. Alonso, D.L. Patrick, L.M. Bouter, H.C. de Vet, The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content, *BMC Med. Res. Methodol.* 10 (2010) 22. <https://doi.org/10.1186/1471-2288-10-22>.
- [32] C.B. Terwee, S.D.M. Bot, M.R. de Boer, D.A.W.M. van der Windt, D.L. Knol, J. Dekker, L.M. Bouter, H.C.W. de Vet, Quality criteria were proposed for measurement properties of health status questionnaires, *J. Clin. Epidemiol.* 60 (2007) 34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>.
- [33] J.C. Nunnally, I. Bernstein, *Psychometric Theory*, 3rd ed., McGraw-Hill, New York, 1994.

- [34] F. Guillemin, C. Bombardier, D. Beaton, Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines, *J. Clin. Epidemiol.* 46 (1993) 1417–1432.
- [35] A. Bellier, P. Chaffanjon, E. Krupat, P. Francois, J. Labarère, Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills, *PloS One.* 15 (2020) e0230672. <https://doi.org/10.1371/journal.pone.0230672>.
- [36] B. Fossli Jensen, P. Gulbrandsen, J.S. Benth, F.A. Dahl, E. Krupat, A. Finset, Interrater reliability for the Four Habits Coding Scheme as part of a randomized controlled trial, *Patient Educ. Couns.* 80 (2010) 405–409. <https://doi.org/10.1016/j.pec.2010.06.032>.
- [37] J. Burt, G. Abel, N. Elmore, J. Campbell, M. Roland, J. Benson, J. Silverman, Assessing communication quality of consultations in primary care: initial reliability of the Global Consultation Rating Scale, based on the Calgary-Cambridge Guide to the Medical Interview, *BMJ Open.* 4 (2014) e004339. <https://doi.org/10.1136/bmjopen-2013-004339>.
- [38] T. Fassaert, S. van Dulmen, F. Schellevis, J. Bensing, Active listening in medical consultations: development of the Active Listening Observation Scale (ALOS-global), *Patient Educ. Couns.* 68 (2007) 258–264. <https://doi.org/10.1016/j.pec.2007.06.011>.
- [39] L.C. Zandbelt, E.M.A. Smets, F.J. Oort, H.C.J.M. de Haes, Coding patient-centred behaviour in the medical encounter, *Soc. Sci. Med.* 1982. 61 (2005) 661–671. <https://doi.org/10.1016/j.socscimed.2004.12.006>.
- [40] I. Enzer, J. Robinson, M. Pearson, S. Barton, T. Walley, A reliability study of an instrument for measuring general practitioner consultation skills: the LIV-MAAS scale, *Int. J. Qual. Health Care J. Int. Soc. Qual. Health Care.* 15 (2003) 407–412.
- [41] J. Robinson, T. Walley, M. Pearson, D. Taylor, S. Barton, Measuring consultation skills in primary care in England: evaluation and development of content of the MAAS scale, *Br. J. Gen. Pract. J. R. Coll. Gen. Pract.* 52 (2002) 889–893.
- [42] I. Scholl, J. Nicolai, S. Pahlke, L. Kriston, E. Krupat, M. Härter, The German version of the Four Habits Coding Scheme - association between physicians' communication and shared decision making skills in the medical encounter, *Patient Educ. Couns.* 94 (2014) 224–229. <https://doi.org/10.1016/j.pec.2013.10.006>.
- [43] E. Krupat, R. Frankel, T. Stein, J. Irish, The Four Habits Coding Scheme: validation of an instrument to assess clinicians' communication behavior, *Patient Educ. Couns.* 62 (2006) 38–45. <https://doi.org/10.1016/j.pec.2005.04.015>.
- [44] T.J. Gallagher, P.J. Hartung, S.W. Gregory, Assessment of a measure of relational communication for doctor-patient interactions, *Patient Educ. Couns.* 45 (2001) 211–218.

- [45] L. Côté, A. Savard, R. Bertrand, [Evaluation of the physician-patient relationship competence. Development and validation of an assessment instrument]., *Can. Fam. Physician*. 47 (2001) 512–518.
- [46] F. Lehmann, L. Côté, A. Bourque, D. Fontaine, Physician-Patient Interaction: A Reliable and Valid Check-list of Quality, *Can. Fam. Physician Med. Fam. Can.* 36 (1990) 1711–1716.
- [47] L. Crocker, J. Algina, *Introduction to Classical and Modern Test Theory.*, Harcourt, New York, 1986.
- [48] D. Wilkin, L. Hallam, M.-A. Doggett, *Measures of need and outcome for primary health care*, Oxford University Press, New York, NY, US, 1992.
- [49] C.B. Terwee, C. a. C. Prinsen, A. Chiarotto, M.J. Westerman, D.L. Patrick, J. Alonso, L.M. Bouter, H.C.W. de Vet, L.B. Mokkink, COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study, *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* 27 (2018) 1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>.
- [50] T.V. Perneger, A. Leplège, J.F. Etter, Cross-cultural adaptation of a psychometric instrument: two methods compared, *J. Clin. Epidemiol.* 52 (1999) 1037–1046.

2.2. **Partie 2)** Adaptation transculturelle et validation des propriétés psychométriques de l'échelle 4-Habits Coding Scheme (4-HCS)

Article publié — Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills

A partir des résultats de la revue systématique précédente, l'instrument de mesure standardisé 4-Habits Coding Scheme (4-HCS) semblait particulièrement pertinent pour l'évaluation des compétences relationnelles dans leur ensemble et dans des contextes variés d'utilisation, c'est-à-dire en simulation dans le cadre de la formation initiale ou avec des patients réels dans le cadre de la formation continue. Cependant, il n'existait pas de version française validée de cet instrument. Ainsi, dans cet article publié le 16 avril 2020 dans la revue PlosOne, nous avons réalisé l'adaptation trans-culturelle de l'échelle, puis nous avons validé les propriétés psychométriques de l'instrument.

Cette étude a été réalisée grâce au soutien financier de la Fondation MACSF.

Contexte

L'échelle 4-Habits Coding Scheme (4-HCS) est un instrument standardisé conçu pour évaluer les compétences relationnelles des médecins du point de vue d'un évaluateur externe, sur la base de consultations vidéo-enregistrées.

Objectif

Nous avons réalisé l'adaptation transculturelle de l'échelle 4-HCS en langue française et nous avons ensuite évalué ses propriétés psychométriques.

Méthodes

L'échelle 4-HCS a fait l'objet d'une adaptation trans-culturelle en effectuant des traductions et rétro-traductions avec des traducteurs professionnels indépendants, conformément aux recommandations internationales. Quatre évaluateurs ont ensuite évalué 200 consultations d'étudiants en médecine vidéo-enregistrées avec des patients standardisés, en utilisant la version française de l'échelle 4-HCS. Nous avons alors réalisé une analyse factorielle, une analyse de la cohérence interne, de la validité de construit et de la fiabilité de l'instrument de mesure.

Résultats

Le score global moyen de l'échelle 4-HCS était de 76,44 (écart-type : 12,34), sans effet plancher ou plafond dans les sous-échelles. La durée médiane de l'évaluation était de 8 min (étendue : 4 à 19). Le coefficient alpha de Cronbach était de 0,94 pour l'échelle 4-HCS globale, allant de 0,72 à 0,88 pour les sous-échelles (quatre dimensions). Dans l'analyse factorielle confirmatoire, les statistiques de qualité de l'ajustement n'ont pas corroboré la structure hypothétique des 4 dimensions (« habitudes »). L'analyse factorielle exploratoire a donné lieu à deux dimensions, avec la fusion de trois dimensions conceptuellement liées en une seule dimension et un recouplement important pour 15 des 23 items. La médiane des estimations du coefficient de corrélation intra-classe d'accord absolu était de 0,74 (étendue : 0,68 à 0,84) et de 0,85 (étendue : 0,76 à 0,91) pour la fiabilité inter- et intra-juges des sous-échelles, respectivement.

Conclusion

La version française de l'échelle 4-HCS démontre une cohérence interne satisfaisante, mais nécessite l'utilisation de deux évaluateurs indépendants pour atteindre une fiabilité acceptable. La structure factorielle sous-jacente de la version originale américaine et les adaptations transculturelles du 4-HCS méritent d'être étudiées de manière plus approfondie dans de futures recherches.

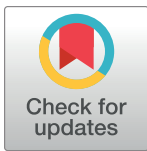
RESEARCH ARTICLE

Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills

Alexandre Bellier^{1,2,3,4*}, Philippe Chaffanjon², Edward Krupat⁵, Patrice Francois¹, José Labarère^{1,2,3,4}

1 Quality of Care Unit, Grenoble Alpes University Hospital, Grenoble, France, **2** School of Medicine, Grenoble Alpes University, Grenoble, France, **3** Computational and Mathematical Biology Team, TIMC-IMAG, UMR 5525, CNRS, Grenoble Alpes University, Grenoble, France, **4** CIC 1406, INSERM, Grenoble Alpes University, Grenoble, France, **5** Center for Evaluation, Harvard Medical School, Harvard University, Boston, Massachusetts, United States of America

* abellier@chu-grenoble.fr



OPEN ACCESS

Citation: Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J (2020) Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills. PLoS ONE 15(4): e0230672. <https://doi.org/10.1371/journal.pone.0230672>

Editor: Gian Mauro Manzoni, eCampus University, ITALY

Received: September 4, 2019

Accepted: March 5, 2020

Published: April 16, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0230672>

Copyright: © 2020 Bellier et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The dataset are available in Harvard Dataverse (view at <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/IZTLWR>).

Abstract

Background

The Four Habits Coding Scheme (4-HCS) is a standardized instrument designed to assess physicians' communication skills from an external rater's perspective, based on video-recorded consultations.

Objective

To perform the cross-cultural adaptation of the 4-HCS into French and to assess its psychometric properties.

Methods

The 4-HCS was cross-culturally adapted by conducting forward and backward translations with independent translators, following international guidelines. Four raters rated 200 video-recorded medical student consultations with standardized patients, using the French version of the 4-HCS. We examined the internal consistency, factor structure, construct validity, and reliability of the 4-HCS.

Results

The mean overall 4-HCS score was 76.44 (standard deviation, 12.34), with no floor or ceiling effects across subscales. The median rating duration of rating was 8 min (range, 4–19). Cronbach's alpha was 0.94 for the overall 4-HCS, ranging from 0.72 to 0.88 across subscales. In confirmatory factor analysis, goodness-of-fit statistics did not corroborate the hypothesized 4-habit structure. Exploratory factor analysis resulted in two dimensions, with the merging of three conceptually related habits into a single dimension and substantial cross-loading for 15 out of 23 items. Median average absolute-agreement intra-class

Funding: This study was supported by unrestricted grants from the MACSF Corporate Foundation (19,951, AB: <https://www.macsf-exerciceprofessionnel.fr/Publications-actions-et-mecenat/Fondation-MACSF>) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist

correlation coefficient estimates were 0.74 (range, 0.68–0.84) and 0.85 (range, 0.76–0.91) for inter- and intra-rater reliability of habit subscales, respectively.

Conclusion

The French version of the 4-HCS demonstrates satisfactory internal consistency but requires the use of two independent raters to achieve acceptable reliability. The underlying factor structure of the original US version and cross-cultural adaptations of the 4-HCS deserve further investigation.

Introduction

Physician communication skills are key components of effective medical consultations [1] and comprise core physician competences that are most desired by patients [2]. Evidence has accumulated, supporting the conclusion that high-quality communication relates with enhanced patient satisfaction [3], greater adherence to treatment [4], better health outcomes [5], and decreased risk of malpractice claims [6]. Many organizations have therefore implemented structured training programs and routinely assessed physicians' communication skills [3, 7].

The Four Habits Coding Scheme (4-HCS) is a standardized instrument designed to assess 23 physician communication skills or behaviors from an external rater's perspective, based on video-recorded consultations [8]. The 4-HCS is based on the conceptual framework of the "Four Habits Model," a training program that was developed within the US Kaiser Permanente Health Maintenance Organization and implemented for teaching effective communication skills to thousands of clinicians in this organization over the two last decades [9]. The Four Habits Model refers to basic medical interview tasks that are organized within four dimensions for didactic purposes, namely, Invest in the beginning (six items), Elicit the patient's perspective (three items), Demonstrate empathy (four items), and Invest in the end (ten items) [9].

The original US version of the 4-HCS demonstrated acceptable inter-rater reliability and evidence for construct validity despite moderate internal consistency, with Cronbach's alpha coefficients ranging from 0.51 to 0.81 across the Four Habits, over 100 video-recorded physician-patient encounters [8]. Since its original development, the 4-HCS has been utilized outside the Kaiser Permanente system [9–13]. Cross-cultural adaptations of the 4-HCS have been published in different languages: Norwegian [14], German [15], and Brazilian Portuguese [16]. Given the potential of the 4-HCS for assessing baseline communication skills and measuring the effectiveness of a training program aiming to alter communication skills [9], there is a need for a French version that can be used with medical students during the 4-year competency-based communication curriculum. To our knowledge, only three studies examined the psychometric properties of the 4-HCS and none has investigated the underlying factor structure [8, 14, 15]. Although the findings of previous studies were promising, the authors recommended examining the validity and reliability of the 4-HCS further in different settings and populations [8].

In the present study, we aimed to perform the cross-cultural adaptation of the 4-HCS into French and to assess the psychometric properties of the adapted version, using the original data of video-recorded medical student consultations with standardized patients.

Materials and methods

Study design

A two-step procedure was used. Firstly, the 4-HCS was translated and cross-culturally adapted into French. Secondly, the psychometric properties of the French version were studied, with regard to internal consistency, validity, and reliability. The present study was conducted in the Grenoble Alpes University Hospital and the University Grenoble Alpes School of Medicine, France.

Cross-cultural adaptation of the 4-HCS into French

The translation and cross-cultural adaptation of the source version of the 4-HCS and its code-book were performed by Mapi Language Services, according to published guidelines [17]. Mapi Language Services is an international company with expertise in the field of translation and cross-cultural adaptation of patient reported outcome measures (www.mapigroup.com/services/language-services/).

The aim of the cross-cultural adaptation process was to obtain a French translation that was conceptually equivalent to the US source version, culturally relevant to the French context, and easily understood by the people who would use the instrument. For this purpose, we used a rigorous methodology involving input from the 4-HCS developer on conceptual issues and a centralized review process coordinated by a consultant with experience in the field. This process included a common understanding of the 4-HCS concepts by all participants involved in the project, quality control by translators, and discussion about translation decisions at each step.

Practically speaking, the 4-HCS developer (EK) was contacted to obtain permission to use and translate the instrument and to invite him to participate in the project. Two qualified native French-speaking translators independently translated the source version of the 4-HCS into French. A single version was obtained after a reconciliation meeting of the two translators. Then this version was back-translated into English by a third qualified native English speaker, who was blinded to the original US version. The back translation was reviewed for semantic and operational equivalence against the source version of the 4-HCS. We followed a universalist approach for equivalence, assessing conceptual, item, semantic, operational, measurement and functional equivalence [18]. After the resulting version was pretested by two raters using 63 video-recorded medical student consultations with standardized patients, minor adjustments were made to obtain the final version. The principal investigator checked the proofs of the final version and corrected any errors.

4-HCS scoring

In accordance with the source version of the 4-HCS, each item was rated on a 5-point Likert scale, ranging from 1 to 5, with higher scores denoting better performance. The midpoint (i.e., 3) and the two endpoints (i.e., 1 and 5) were anchored, with specific behavioral descriptions [8]. The raters were encouraged to use the midpoint and endpoint categories, with other categories (i.e., 2 and 4) to be used only if they thought that communication skills fell between the anchored points [8]. This approach ensured full use of the 5-point Likert scale [8].

An overall communication skill score was computed by summing ratings for the individual items, ranging from 23 (i.e., less effective) to 115 (i.e., more effective). Four subscale scores were also computed, each of them corresponding to a key dimension of communication skills (i.e., Invest in the beginning [range, 6–30], Elicit the patient's perspective [range, 3–15], Demonstrate empathy [range, 4–20], and Invest in the end [range, 10–50]).

Physician–patient relationship competence assessment scale

The physician–patient relationship competence assessment scale was developed in French and validated in Canada [19]. This scale consists in 15 items exploring two dimensions of interpersonal skills, namely “understanding of the patient’s disease experience” (eight items) and “efficient and respectful communication” (seven items). Each item could be rated on a 4-point Likert scale, ranging from 15 (i.e., less effective) to 60 (i.e., more effective).

Study sample and data collection

The study sample consisted of video-recorded medical student consultations with standardized patients. All 218 medical students were invited to participate as part of the 4-year competency-based communication curriculum. They were allocated in consecutive alphabetical order based on their surname to the 1st or 2nd semester sessions that took place in October 2017 and April 2018, respectively.

Standardized patients were 20 actors recruited at the Department of Performing Arts in Grenoble Alpes University. Seven standardized medical consultation scenarios were developed and combined with nine personality types or character traits. The actors were instructed on each medical consultation scenario by two clinicians during 2-h sessions. Then they were trained by their improvisation instructor on each personality type or character trait. The combinations of medical consultation scenarios with personality type or character trait were pre-tested as part of a pilot study.

All medical consultations with standardized patients were video-recorded using professional video equipment. Video-recording of consultations made it possible to overcome some of the challenges of direct observation [20]. Video-recording accurately recorded all events that occurred during consultations, allowing raters to verify their observations as many times as necessary [20]. Video-recordings could be rated by different observers without consultations being disrupted [21]. Finally, video-recording allowed for providing medical students with feedback on their own performance [20, 22].

Four raters were recruited for the project, including a full professor of medicine (PC) and a resident in medicine (AB), both with experience in teaching communication skills, and two medical students. They underwent a training session, which consisted in independently coding five video-recorded consultations with the 4-HCS and subsequently discussing these ratings. After completing the training, the raters independently rated consultations, with each consultation being rated by at least three different raters. Additionally, two of the raters evaluated each video-recorded medical student consultation twice, with the two ratings 2 months apart, in order to quantify intra-rater reliability.

Data for both the cross-cultural adaptation of the 4-HCS and the physician–patient relationship competence assessment scale were captured using an electronic case report form. The completion of each item was mandatory, so there were no missing values.

Sample size

A sample size of 200 video-recorded consultations was required for confirmatory factor analysis, based on previous simulations [23]. Assuming a Cronbach alpha coefficient point estimate close to 0.80, this sample size would provide a precision of ± 0.07 (i.e., 95% confidence interval [CI] ranging from 0.73 to 0.87).

Assuming an intra-rater correlation coefficient point estimate of 0.80, we estimated that a sample of 117 video-recorded consultations would achieve 80% power to demonstrate that it would be higher than 0.70, with a 0.05 two-sided significance level [24]. Assuming an inter-rater correlation coefficient point estimate of 0.80 with four raters, we estimated that a sample

of 68 video-recorded consultations would achieve 80% power to demonstrate that it would be higher than 0.70, with a 0.05 two-sided significance level [24].

Statistical analysis

Descriptive statistics. To account for inter-rater variability, we computed the students' average 4-HCS overall and subscale rating scores [25]. The 4-HCS overall and subscale scores were reported as means along with standard deviations. The numbers and percentages of responses on anchor points for items and overall and subscale scores were examined to detect floor or ceiling effects. Floor and ceiling effects lower than 15% for subscale scores were considered acceptable [26].

Internal consistency. Internal consistency was evaluated through average inter-item correlation, item-rest correlation (i.e., the correlation between an item and the score that was formed by all other items in the subscale), and Cronbach's alpha coefficient [27]. The internal consistency criterion was fulfilled for item-rest correlation >0.40 , and Cronbach's alpha >0.70 was considered satisfactory [26].

Internal structure. The internal structure of the cross-cultural adaptation of the 4-HCS into French was verified using confirmatory and exploratory factor analysis, following current guidelines [27]. First, structural equation modeling of the four predefined habits was carried out with the 23 items assigned to the intended habits to determine whether the video-recorded medical consultation data fit with the internal structure of the original US version of the 4-HCS. This structural equation modeling corresponded to an external model representing the relationships between latent variables (i.e., the four predefined habits) and the manifest variables (i.e., the 23 related items). Various goodness-of-fit statistics were obtained, including the comparative fit index (CFI) [28], the standardized root mean squared residual (SRMR), and the root mean square error of approximation (RMSEA) along with its 90% CI [29]. A CFI value of 0.90 or higher and a SRMR value lower than 0.08 were considered indicative of satisfactory model fit. A 90% CI lower bound for the RMSEA estimate lower than 0.05 would not reject the hypothesis that the fit was close. An upper bound higher than 0.10 would not reject the hypothesis that the fit was poor.

Second, exploratory factor analysis was performed in order to examine possible alternative structures to the original US version of the 4-HCS. An orthogonal rotation method (Varimax) of factors with eigenvalues higher than 1.00 was used, assuming that they were independent [30]. Primary loadings on intended dimensions higher than 0.40 with cross-loadings lower than 0.30 were considered satisfactory.

Construct validity. We assessed construct validity by comparing 4-HCS overall and subscale score values between first and second-semester video-recorded medical consultations. We hypothesized that mean scores were higher for medical consultations recorded during the second semester. Indeed, second semester students were assumed to be more experienced in basic medical interview tasks and more sensitized to communication skills than their counterparts evaluated during the first semester. We also examined convergent validity between the 4-HCS and the physician-patient relationship competence assessment scale, using Pearson correlation coefficients.

Reliability. Inter- and intra-rater reliability assessment complied with the *Guidelines for Reporting Reliability and Agreement Studies* (GRAAS) [31]. The reliability of the French version of the 4-HCS overall and subscale scores was quantified by the intra-class correlation coefficient (ICC) [32]. The ICC is suitable for reliability studies with unbalanced designs involving different sets of raters [33]. Both absolute- and consistency-of-agreement ICCs were computed. Under the absolute-agreement approach, the ratings were considered in absolute

agreement if the 4-HCS scores from all raters matched exactly [32]. Under the consistency-of-agreement approach, the ratings were considered consistent if the 4-HCS scores from any two raters differed by the same constant value for all video-recorded consultations. This implied that raters gave the same ranking to all video-recorded consultations [34]. Individual and average ICCs were estimated, with average ICCs computed over two raters. Although the agreement measured between individual ratings is more common, the use of average ICCs was indicated in this study because the 4-HCS was intended to be used by teams of raters for assessing video-recorded medical consultations. An ICC value equal to or higher than 0.70 was indicative of satisfactory reliability [14].

The study protocol was approved by the Comité d’Ethique du Centre d’Investigation Clinique de Clermont-Ferrand, France (IRB 5891). All participants received information from the principal investigator about the study’s overall purpose and the confidentiality requirements and they then provided written informed consent.

Results

Mapi Language Services translated the 4-HCS scale in January 2018 and issued a translation validation certificate on February 19, 2018. They scrupulously complied with the protocol drawn up according to international standards, producing a cross-cultural adaptation of the 4-HCS scale in line with expectations. The French version of the 4-HCS is shown in [S1 Appendix](#). The full version included a translation of the “codebook,” a detailed scoring guide for each item with examples of behaviors and suggested ratings.

Of 218 eligible medical students, 200 (92%) participated in the study. A total of 200 consultations with standardized patients were video-recorded, including 115 and 85 during the first and second semesters, respectively ([Fig 1](#)). The median duration was 8 min (range, 4–19 min) per video-recorded consultation. Inter-rater reliability assessment involved 800 ratings while 400 ratings contributed to intra-rater reliability assessment ([Fig 1](#)).

The mean 4-HCS score was 76.44 (standard deviation, 12.34) for 200 video-recorded medical consultations, with no floor or ceiling effects observed for subscales ([Table 1](#)). Yet the highest (5/5) and lowest values (1/5) accrued more than 15% of the respondents for three items and one item, respectively. Mean 4-HCS scores ranged from 66.97 (SD, 10.26) based on 200 video-recorded consultations for rater 1 (i.e., the least experienced rater) to 93.42 (SD 13.46) based on 85 video-recorded consultations for rater 4 (i.e., the most experienced rater).

Cronbach’s alpha was 0.94 (95% CI, 0.93–0.95) for the 4-HCS, ranging from 0.72 to 0.88 across habit subscales ([Table 1](#)). All but one item fulfilled the internal consistency criterion, with item-rest correlations higher than 0.40. The exception was item 16 with the item-rest correlation as low as 0.15. Removing this item from the “Invest in the end” habit subscale improved Cronbach’s alpha from 0.88 to 0.90.

In confirmatory factor analysis, structural equation modeling of 23 items apportioned in four latent factors yielded CFI (0.79) and SRMR (0.09) estimates that did not achieve recommended thresholds (CFI >0.90 and SRMR <0.08, respectively). The RMSEA estimate was 0.12 (90% CI, 0.10–0.13), with the 90% CI lower bound not rejecting the hypothesis that the fit was close and the 90% CI upper bound not rejecting the hypothesis that the fit was poor.

Exploratory factor analysis of the 23 items identified four principal components with eigenvalues higher than 1.0 and explaining 66.8% of overall variance ([Table 2](#)). Graphical assessment of the scree plot suggested that the instrument was close to unidimensionality ([S2 Appendix](#)), supporting the use of an overall 4-HCS score. Yet, the factorial structure for the French version of the 4-HCS departed from the hypothesized four-dimension structure. Twenty-one out of 23 items had primary factor loadings over 0.40 while 15 items yielded

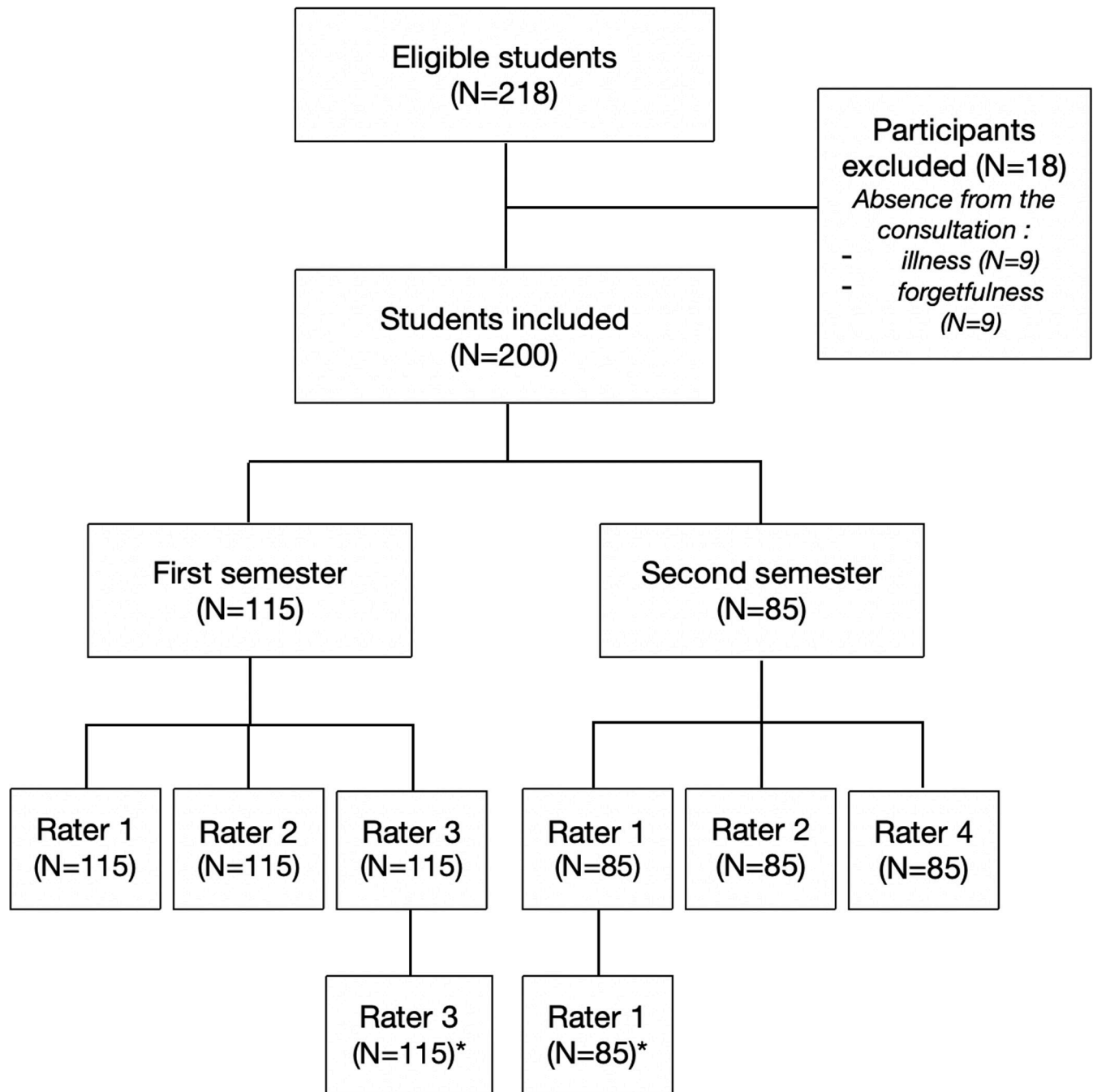


Fig 1. Flow of medical students and raters throughout the study.

<https://doi.org/10.1371/journal.pone.0230672.g001>

cross-loadings over 0.30. Exploratory factor analysis of the French version of the 4-HCS resulted in the merging of three habits (namely, Invest in the beginning [all six items], Elicit the patient's perspective [two out of three items], and Demonstrate empathy [all four items]) in a single dimension.

As hypothesized, all mean habit scores were significantly higher for medical student consultations recorded at the second semester (Table 3). Most pairwise Pearson correlation

Table 1. Summary statistics and internal consistency for the cross-cultural adaptation of the 4 Habit Coding Scheme into French (n = 200).

Habit [range]-Item	Mean score (SD)		Ceiling effect, n (%)		Floor effect, n (%)		Average inter-item correlation	Item-total correlation	Item-rest correlation*	Cronbach Alpha†
1. Invest in the beginning [6–30]	18.77	(2.80)	0	(0)	0	(0)	0.48	-	-	0.80
1. Show familiarity	3.09	(0.18)	0	(0)	0	(0)	0.43	0.55	0.50	0.81
2. Greet warmly	3.18	(0.27)	0	(0)	0	(0)	0.48	0.64	0.57	0.79
3. Engage in small talk	2.66	(0.89)	6	(3.0)	4	(2.0)	0.46	0.79	0.62	0.76
4. Question style	3.25	(0.74)	2	(1.0)	0	(0)	0.52	0.84	0.72	0.73
5. Expansion of concerns	3.51	(0.70)	0	(0)	4	(2.0)	0.50	0.82	0.70	0.73
6. Elicit full agenda	3.07	(0.81)	3	(1.5)	1	(0.5)	0.47	0.77	0.60	0.76
2. Elicit the patient's perspective [3–15]	10.01	(2.06)	0	(0)	0	(0)	0.48	-	-	0.72
7. Patient's understanding	3.70	(0.78)	0	(0)	4	(2.0)	0.48	0.79	0.55	0.62
8. Goals for visit	2.81	(0.79)	6	(3.0)	0	(0)	0.51	0.82	0.60	0.57
9. Impact on life	3.49	(0.98)	2	(1.0)	14	(7.0)	0.43	0.81	0.49	0.72
3. Demonstrate empathy [4–20]	13.75	(2.95)	0	(0)	3	(1.5)	0.52	-	-	0.87
10. Encourage emotional expression	3.33	(0.90)	1	(0.5)	9	(4.5)	0.61	0.90	0.84	0.81
11. Accept feelings	3.63	(0.73)	1	(0.5)	4	(2.0)	0.61	0.87	0.81	0.82
12. Identify feelings	2.60	(1.04)	26	(13.0)	23	(1.5)	0.55	0.84	0.73	0.84
13. Show good nonverbal behavior	4.19	(0.71)	0	(0)	42	(20.5)	0.53	0.76	0.66	0.84
4. Invest in the end [10–50]	33.91	(6.13)	0	(0)	0	(0)	0.46	-	-	0.88
14. Use patient's frame of reference	3.60	(0.81)	0	(0)	15	(7.5)	0.45	0.70	0.56	0.86
15. Allow time to absorb	4.37	(0.56)	0	(0)	33	(16.5)	0.36	0.55	0.43	0.88
16. Give clear explanation	4.24	(0.76)	0	(0)	63	(31.5)	0.12	0.29	0.15	0.90
17. Offer rationale for tests	3.56	(0.95)	8	(4.0)	10	(5.0)	0.45	0.73	0.63	0.86
18. Test for comprehension	3.33	(0.96)	8	(4.0)	4	(2.0)	0.59	0.89	0.84	0.84
19. Involve in decision	3.01	(0.83)	7	(3.5)	2	(1.0)	0.57	0.86	0.81	0.84
20. Explore plan acceptability	3.21	(0.98)	6	(3.0)	3	(1.5)	0.58	0.87	0.82	0.84
21. Explore barriers	2.02	(0.84)	47	(23.5)	0	(0)	0.46	0.73	0.64	0.86
22. Encourage questions	3.56	(0.93)	1	(0.5)	25	(12.5)	0.44	0.72	0.61	0.86
23. Plan for follow-up	3.00	(0.95)	8	(4.0)	1	(0.5)	0.43	0.70	0.58	0.87
Overall [23–115]	76.44	(12.34)					0.42	-	-	0.94

Abbreviations: SD, standard deviation.

* Item-rest correlation was computed as the correlation between an item and the composite score that was formed by all other items in the habit.

† Item Cronbach alpha was computed for composite score that was formed by all other items in the habit.

<https://doi.org/10.1371/journal.pone.0230672.t001>

coefficients between habit and physician-patient relationship competence assessment scale scores were higher than 0.70, indicating satisfactory convergent validity (Table 4).

None of the four habits fulfilled the 0.70 inter-rater reliability criterion, with individual absolute-agreement ICC point estimates ranging from 0.42 to 0.64 (Table 5). Interestingly, the average absolute-agreement ICC for the 4-HCS and three out of four habits fulfilled the inter-rater reliability criterion of 0.70.

The individual absolute-agreement ICC was 0.72 for intra-rater reliability of the 4-HCS score, with point estimates ranging from 0.45 to 0.71 across habits. The average absolute-agreement ICC for the 4-HCS and all habits fulfilled the 0.70 intra-rater reliability criterion.

Table 2. Exploratory factor analysis for the 23 items of the cross-cultural adaptation of the 4-Habit Coding Scheme into French after orthogonal Varimax rotation (n = 200)*.

Habit-Item	Factor 1	Factor 2	Factor 3	Factor 4
1. Invest in the beginning				
1. Show familiarity	.41			.40
2. Greet warmly	.40	.34		.33
3. Engage in small talk	.58			
4. Question style	.58	.49	.32	
5. Expansion of concerns	.74	.33		
6. Elicit full agenda	.64	.32		
2. Elicit the patient's perspective				
7. Patient's understanding	.53	.61		
8. Goals for visit	.48	.45	.47	.35
9. Impact on life	.58			
3. Demonstrate empathy				
10. Encourage emotional expression	.91			
11. Accept feelings	.74	0.35		
12. Identify feelings	.86			
13. Show good nonverbal behavior	.54	.37		
4. Invest in the end				
14. Use patient's frame of reference	.33	.73		
15. Allow time to absorb	.36			
16. Give clear explanation				
17. Offer rationale for tests		.71		
18. Test for comprehension	.31	.84	.42	
19. Involve in decision	.39	.76	.67	
20. Explore plan acceptability		.78	.63	
21. Explore barriers		.61	.46	
22. Encourage questions		.69		
23. Plan for follow-up		.56	.47	
Overall variance explained, %	46.5	8.8	6.7	4.8

* Values are item loadings $\geq .30$

<https://doi.org/10.1371/journal.pone.0230672.t002>

After restricting our analytical sample to the most experienced raters (raters #1 and #2), the individual absolute-agreement ICC was 0.83 for inter-rater reliability and 0.89 for intra-rater reliability.

Table 3. Comparison of 4-Habit Coding Scheme scores for medical student consultations recorded during first and second semesters.

Habit [range]	1 st semester (n = 115)		2 nd semester (n = 85)		P
1. Invest in the beginning [6–30]	17.54	(2.56)	19.83	(2.56)	< .001
2. Elicit the patient's perspective [3–15]	8.88	(1.93)	10.98	(1.63)	< .001
3. Demonstrate empathy [4–20]	13.02	(3.34)	14.36	(2.43)	< .001
4. Invest in the end [10–50]	31.84	(5.95)	31.84	(35.70)	< .001
Overall [23–115]	71.28	(12.09)	80.87	(10.77)	< .001

* Values are mean (standard deviation)

<https://doi.org/10.1371/journal.pone.0230672.t003>

Table 4. Correlation of the cross-cultural adaptation of the 4-Habit Coding Scheme into French and the physician–patient relationship competence assessment scale (n = 200)*.

Habit [range]	Physician–patient relationship competence assessment scale		
	1. Understanding of the patient’s experience	2. Communication, consultation management	Overall
1. Invest in the beginning	.85	.61	.81
2. Elicit the patient’s perspective	.92	.47	.79
3. Demonstrate empathy	.77	.64	.76
4. Invest in the end	.70	.79	.78
Overall	.90	.74	.90

* Values are Pearson correlation coefficients. All *P*-values were < .001.

<https://doi.org/10.1371/journal.pone.0230672.t004>

Discussion

The use of validated standardized instruments is advocated for assessing physicians’ interpersonal skills [8]. Indeed, validated instruments are likely to accurately reflect the concept to be measured while standardized instruments allow large-scale comparisons of physicians’ communication skills across studies [30, 35]. The 4-HCS was therefore developed and validated from over 1,025 video-recorded medical consultations across various settings in the US and Western Europe (Table 6). Yet, our study was the first to examine the underlying factor structure of the 4-HCS and to report on its cross-cultural adaptation into French.

Each of the 200 video-recorded medical consultations in this study was rated without missing data by three out of four different raters, reflecting the acceptability and feasibility of the French version of the 4-HCS. Although the consultations were simulated with standardized patients and involved 4-year medical students in this study, the mean overall and subscale 4-HCS scores were consistent with previous reports (Table 6).

Surprisingly, the French version of the 4-HCS yields better performance regarding internal consistency than the original US instrument and previous cross-cultural adaptations. Cronbach’s alpha was higher than 0.70 across habit subscales (median 0.83; range, 0.72–0.88) and compared favorably with those reported in the original US development (median, 0.66; range, 0.51–0.81) and German cross-cultural adaptation (median, 0.39; range, 0.31–0.46) (Table 6). This finding contrasts with previous studies that usually report worse performance for cross-cultural adaptations in comparison with original standardized instruments [30]. Krupat et al. were not concerned by the moderate internal consistency of the original version of the 4-HCS and speculated that successful training in communication would result in more effective communication and therefore greater internal consistency [8]. Another potential explanation for this discrepancy may be that the conceptual framework of the 4-HCS lacks generalizability or

Table 5. Absolute-agreement intra-class correlation coefficient estimates for inter- and intra-rater reliability for the cross-cultural adaptation of the 4-Habit Coding Scheme into French (n = 200).

Habit	Inter-rater ICC (95% CI)				Intra-rater ICC (95% CI)			
	Individual		Average		Individual		Average	
1. Invest in the beginning	.45	(.36 to .53)	.71	(.63 to .77)	.59	(.52 to .65)	.85	(.81 to .88)
2. Elicit the patient’s perspective	.42	(.33 to .60)	.68	(.60 to .75)	.45	(.37 to .52)	.76	(.71 to .81)
3. Demonstrate empathy	.53	(.45 to .60)	.77	(.71 to .82)	.58	(.51 to .64)	.85	(.81 to .88)
4. Invest in the end	.64	(.58 to .71)	.84	(.80 to .88)	.71	(.66 to .76)	.91	(.89 to .93)
Overall	.60	(.53 to .67)	.82	(.77 to .86)	.72	(.67 to .77)	.91	(.89 to .93)

Abbreviations: CI, confidence interval; ICC, intra-class correlation.

<https://doi.org/10.1371/journal.pone.0230672.t005>

Table 6. Primary studies reporting on the development or cross-cultural adaptation of the 4-Habit Coding Scheme.

Author, year	Krupat, 2006	Fossli Jensen, 2010	Clayton, 2011	Scholl, 2014	Present study
Country	USA	Norway	USA	Germany	France
Setting	Hospital	Hospital	Family practice clinics	Primary and specialty consultations	School of medicine
Recruitment period	1994	2007–2008	-	2009–2010	2017–2018
Experience	Resident and senior staff	Resident and senior staff	Resident and senior staff	Senior staff	Medical students
No. physicians	50	71	21	22	200
Simulated consultations	No	No	No	No	Yes
Recording	Video-recorded	Video-recorded	Video-recorded	Audio-taped	Video-recorded
No. consultations	100	497	174	54	200
Mean score					
1. Invest in the beginning	17.7	-	24.1	12.1	18.8
2. Elicit the patient's perspective	7.6	-	11.5	4.5	10.0
3. Demonstrate empathy	11.3	-	14.5	-	13.7
4. Invest in the end	31.5	-	33.0	26.5	33.9
Overall	68.0	60.1	83.1	-	76.4
Exploratory factor analysis	Not performed	Not performed	Not performed	Not performed	4 principal components (66.8% of overall variance)
Cronbach's alpha					
1. Invest in the beginning	.71	-	-	.41	.80
2. Elicit the patient's perspective	.51	-	-	.46	.72
3. Demonstrate empathy	.81	-	-	.38	.87
4. Invest in the end	.61	-	-	.31	.88
Overall	-	.85	-	-	.94
Inter-rater reliability*					
1. Invest in the beginning	.70	-	.48	.83	.45
2. Elicit the patient's perspective	.80	-	.57	.79	.42
3. Demonstrate empathy	.71	-	.39	.85	.53
4. Invest in the end	.69	-	.65	.78	.64
Overall	.72	.78	.72	-	.60
Intra-rater reliability*					
1. Invest in the beginning	-	-	-	.87	.59
2. Elicit the patient's perspective	-	-	-	.72	.45
3. Demonstrate empathy	-	-	-	.84	.58
4. Invest in the end	-	-	-	.83	.71
Overall	-	-	-	-	.72
Author, year	Krupat, 2006	Fossli Jensen, 2010	Clayton, 2011	Scholl, 2014	Present study
2. Elicit the patient's perspective	-	-	-	.72	.45
3. Demonstrate empathy	-	-	-	.84	.58
4. Invest in the end	-	-	-	.83	.71
Overall	-	-	-	-	.72

* Inter-rater reliability was quantified by the Pearson correlation coefficient in the studies by Krupat et al. and Clayton et al., and by the intra-class correlation coefficient in the study by Fossli Jensen et al.

‡ Inter- and intra-rater reliability scores were quantified by computing absolute agreement intra-class correlation coefficient in the study by Scholl et al. In the present study, individual absolute agreement intra-class correlation coefficient was used for assessing inter- and intra-rater reliability.

<https://doi.org/10.1371/journal.pone.0230672.t006>

robustness and therefore does not apply equally to all to target populations' experience. Noticeably, the French version of the 4-HCS was used for rating medical student consultations

with standardized patients while the original US version was used for rating resident and senior staff medical consultations (Table 6).

To our knowledge, no previous study examined the factor structure of the 4-HCS. In confirmatory factor analysis, goodness-of-fit statistics did not support the hypothesized 4-dimension structure for the French version of the 4-HCS. Exploratory factor analysis resulted in two dimensions, with the merging of three conceptually related habits (Invest in the beginning, Elicit the patient's perspective, Demonstrate empathy) in a single dimension. Additionally, substantial cross-loading was observed for 15 out of 23 items, suggesting that the underlying factor structure of the French version of the 4-HCS was questionable.

These findings do not necessarily invalidate the postulated structure of the 4-HCS. Indeed, factor analysis can only discriminate uncorrelated constructs in a data set [36]. That three habits were correlated with each other in the present study sample does not imply that these scales measure the same concept [36]. The medical students participating in the current study were not trained with the Four Habit Model before their communication skills were assessed using the 4-HCS. This might explain why the factor structure for the French version departed from the postulated four-dimension structure of the 4-HCS. Yet, we cannot exclude that this issue is inherent to the original 4-HCS rather than being specific to our study sample. No previous study (including the original development study) examined the factor structure of the 4-HCS. Hence, there is a need for further investigation of the underlying factor structure of the original US version of the 4-HCS.

One item (16. *Give clear explanation*) was not allocated to any empirical dimension in exploratory factorial analysis and also yielded the lowest item-rest correlation, deteriorating the internal consistency of the corresponding habit scale. Altogether, these two observations question the relevance of this item, and therefore its removal from the French version of the 4-HCS should be discussed.

Evidence for convergent validity was provided by the expected correlation between the 4-HCS and physician-patient relationship competence assessment scale scores. Our observation that medical students more experienced in basic interview tasks yielded higher scores for all four habits supported the construct validity of the 4-HCS.

Comparisons of inter- and intra-reliability estimates were confounded by between-study heterogeneity in the types of correlation coefficients used. Only two studies, including the present one, used absolute-agreement ICC (Table 6). ICC is a recommended alternative to Pearson's r coefficient correlation for assessing inter- or intra-reliability [14, 31]. The median absolute-agreement ICC for inter-rater (0.49, range, 0.42–0.64) and intra-rater (0.58, range, 0.45–0.71) reliability for the French version of the 4-HCS were lower than those reported in the German cross-cultural adaptation study (0.81, range, 0.78–0.85 and 0.83, range, 0.72–0.87, respectively). Lower inter-rater reliability might be explained by the use of simulated consultations with standardized patients, the use of video- rather than audio-taped consultations, or varying levels of rater experience in our study. We observed that the greater the level of rater experience, the higher the 4-HCS scores in our study. Inconsistent associations have been reported between experience or seniority and communication scores in objective structured clinical examinations [37]. After restricting our analytical sample to the most experienced raters (i.e., raters #1 and #2), ICC estimates were higher than 0.80 for inter-rater reliability.

Interestingly, substantial improvement in inter-rater (median, 0.74; range, 0.68–0.84) and intra-rater (0.85, range, 0.76–0.91) reliability of habit scores was achieved using average absolute-agreement ICC estimates (Table 5). This latter finding supports the need for using two independent raters to rate communication skills with the 4-HCS based on video-recorded medical consultations, in routine practice.

This study has potential implications for routine assessment of physician communication skills using the cross-cultural adaptation of the 4-HCS into French. First, this study provides evidence on the validity of the 4-HCS scale for simulated consultations with standardized patients. Second, the French version of the 4-HCS scale demonstrated internal consistency, which was even higher than the original US version, allowing international comparisons. Third, our study questioned the hypothesized underlying factor structure of the 4-HCS. Since the 4-HCS was originally developed in the US, the conceptual framework and factor structure may lack generalizability or robustness and therefore may not apply equally well to other countries. Investigating the factor structure of the original US version is required to address this issue. Fourth, the moderate reliability of the French version of the 4-HCS implies that communication skills should be assessed by two independent experienced raters.

This cross-cultural adaptation study has a few caveats that must be considered. First, real patient encounters would have been preferable to standardized patient encounters for assessing communication skills, because of their authenticity [38]. Indeed, simulated consultations with standardized patients differ from real patient consultations in many ways [39]. Simulated patients are not suffering from illness and only attempt to portray the same through their acting. Moreover, recruiting and training standardized patients is time consuming in order to produce a high-quality simulation [40]. Although our study did not explore real patient encounters, standardized patients allowed us to provide a large number of students with reproducible and consistent clinical scenarios of the same level of difficulty [41].

Second, our study was conducted with 4th-year medical students at a single university-affiliated hospital and the findings may not apply to other settings. A broader spectrum of participants would strengthen the confidence in the psychometric properties of this cross-cultural adaptation of the 4-HCS into French.

Third, the 4-HCS and physician–patient relationship competence assessment scale were completed by the same rater so that a halo effect cannot be excluded when assessing convergent validity of the two instruments. Demonstrating that 4-HCS scores correlate with (standardized) patient-reported experience of physician communication skills would provide stronger evidence of construct validity.

Fourth, the one-semester interval separating the two groups of students might be too short to assess the relationship between 4-HCS scores and experience in communication and basic interview tasks, although the associations were significant. The ability of the 4-HCS to discriminate subjects with varying levels of communication skills warrants further investigation. The sensitivity of the French version of the 4-HCS to changes in communication skills also remains to be documented by way of longitudinal studies.

Conclusions

The French version of the 4-HCS demonstrates satisfactory internal consistency but moderate reliability, requiring the use of two independent raters to assess communication skills of medical students based on video-recorded consultations with standardized patients. The empirical factor structure of the French version does not conform with the hypothesized habits of the original 4-HCS. Whether this reflects a specific issue with our cross-cultural adaptation study sample or a more general problem with the instrument is unclear and deserves further investigation.

Supporting information

S1 Appendix. Cross-cultural adaptation of the 4-HCS scale into French.
(DOCX)

S2 Appendix. Scree plot of the 23 items of the 4-HCS scale. (TIF)

Acknowledgments

The authors are indebted to Pauline Bouchet and Ariane Martinez, Grenoble Alpes University, for their collaboration in setting up simulated medical consultations with standardized patients, Noémie Kaladzé and Alexis Dechosal, Grenoble Alpes University, for rating video-recorded medical consultations. The authors thank Alison Foote, Grenoble Alpes University Hospital, and Linda Northrup for their help in editing the manuscript in English.

Author Contributions

Conceptualization: Alexandre Bellier, Philippe Chaffanjon, Edward Krupat, Patrice Francois, José Labarère.

Data curation: Alexandre Bellier.

Formal analysis: Alexandre Bellier, José Labarère.

Funding acquisition: Alexandre Bellier.

Investigation: Alexandre Bellier, Philippe Chaffanjon.

Methodology: Alexandre Bellier, José Labarère.

Project administration: Alexandre Bellier, Philippe Chaffanjon, Patrice Francois, José Labarère.

Resources: Edward Krupat.

Software: Alexandre Bellier, José Labarère.

Supervision: Philippe Chaffanjon, Patrice Francois, José Labarère.

Validation: Philippe Chaffanjon, José Labarère.

Visualization: Philippe Chaffanjon, Patrice Francois.

Writing – original draft: Alexandre Bellier.

Writing – review & editing: Philippe Chaffanjon, José Labarère.

References

1. Ong LML, de Haes JCJM, Hoos AM, Lammes FB. Doctor-patient communication: A review of the literature. *Soc Sci Med*. 1995; 40: 903–918. [https://doi.org/10.1016/0277-9536\(94\)00155-m](https://doi.org/10.1016/0277-9536(94)00155-m) PMID: 7792630
2. Simpson SA, Shaw C, McNamara R. What is the most effective way to maintain weight loss in adults? *BMJ*. 2011; 343: d8042. <https://doi.org/10.1136/bmj.d8042> PMID: 22205707
3. Boissy A, Windover AK, Bokar D, Karafa M, Neuendorf K, Frankel RM, et al. Communication Skills Training for Physicians Improves Patient Satisfaction. *J Gen Intern Med*. 2016; 31: 755–761. <https://doi.org/10.1007/s11606-016-3597-2> PMID: 26921153
4. Zolnieriek KBH, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care*. 2009; 47: 826–834. <https://doi.org/10.1097/MLR.0b013e31819a5acc> PMID: 19584762
5. Kelley JM, Kraft-Todd G, Schapira L, Kossowsky J, Riess H. The influence of the patient-clinician relationship on healthcare outcomes: a systematic review and meta-analysis of randomized controlled trials. *PloS One*. 2014; 9: e94207. <https://doi.org/10.1371/journal.pone.0094207> PMID: 24718585

6. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician-patient communication. The relationship with malpractice claims among primary care physicians and surgeons. *JAMA*. 1997; 277: 553–559. <https://doi.org/10.1001/jama.277.7.553> PMID: 9032162
7. Brown JB, Boles M, Mullooly JP, Levinson W. Effect of clinician communication skills training on patient satisfaction. A randomized, controlled trial. *Ann Intern Med*. 1999; 131: 822–829. <https://doi.org/10.7326/0003-4819-131-11-199912070-00004> PMID: 10610626
8. Krupat E, Frankel R, Stein T, Irish J. The Four Habits Coding Scheme: validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns*. 2006; 62: 38–45. <https://doi.org/10.1016/j.pec.2005.04.015> PMID: 15964736
9. Stein T, Frankel RM, Krupat E. Enhancing clinician communication skills in a large healthcare organization: a longitudinal case study. *Patient Educ Couns*. 2005; 58: 4–12. <https://doi.org/10.1016/j.pec.2005.01.014> PMID: 15950831
10. Fossli Jensen B, Gulbrandsen P, Dahl FA, Krupat E, Frankel RM, Finset A. Effectiveness of a short course in clinical communication skills for hospital doctors: results of a crossover randomized controlled trial (ISRCTN22153332). *Patient Educ Couns*. 2011; 84: 163–169. <https://doi.org/10.1016/j.pec.2010.08.028> PMID: 21050695
11. Blanch-Hartigan D, Hall JA, Roter DL, Frankel RM. Gender bias in patients' perceptions of patient-centered behaviors. *Patient Educ Couns*. 2010; 80: 315–320. <https://doi.org/10.1016/j.pec.2010.06.014> PMID: 20638813
12. Gulbrandsen P, Madsen HB, Benth JS, Laerum E. Health care providers communicate less well with patients with chronic low back pain—a study of encounters at a back pain clinic in Denmark. *Pain*. 2010; 150: 458–461. <https://doi.org/10.1016/j.pain.2010.05.024> PMID: 20705213
13. Clayton MF, Latimer S, Dunn TW, Haas L. Assessing patient-centered communication in a family practice setting: how do we measure it, and whose opinion matters? *Patient Educ Couns*. 2011; 84: 294–302. <https://doi.org/10.1016/j.pec.2011.05.027> PMID: 21733653
14. Fossli Jensen B, Gulbrandsen P, Benth JS, Dahl FA, Krupat E, Finset A. Interrater reliability for the Four Habits Coding Scheme as part of a randomized controlled trial. *Patient Educ Couns*. 2010; 80: 405–409. <https://doi.org/10.1016/j.pec.2010.06.032> PMID: 20685066
15. Scholl I, Nicolai J, Pahlke S, Kriston L, Krupat E, Härter M. The German version of the Four Habits Coding Scheme—association between physicians' communication and shared decision making skills in the medical encounter. *Patient Educ Couns*. 2014; 94: 224–229. <https://doi.org/10.1016/j.pec.2013.10.006> PMID: 24286733
16. Catani RR, Valadares EDS, Lacombe JB, Mendonça TM da S, Silva CHM da, Paro HBM da S. Cross-cultural adaptation of the Four Habits Coding Scheme (4HCS) for teaching and assessing patient-centered communication skills in Brazil. *Cad Saude Publica*. 2018; 34: e00013918. <https://doi.org/10.1590/0102-311X00013918> PMID: 30427408
17. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol*. 2015; 68: 435–441. <https://doi.org/10.1016/j.jclinepi.2014.11.021> PMID: 25698408
18. Harachi TW, Choi Y, Abbott RD, Catalano RF, Bliesner SL. Examining Equivalence of Concepts and Measures in Diverse Samples. *Prev Sci*. 2006; 7: 359–368. <https://doi.org/10.1007/s11121-006-0039-0> PMID: 16845592
19. Côté L, Savard A, Bertrand R. [Evaluation of the physician-patient relationship competence. Development and validation of an assessment instrument]. *Can Fam Physician*. 2001; 47: 512–518. PMID: 11281084
20. Asan O, Montague E. Using video-based observation research methods in primary care health encounters to evaluate complex interactions. *Inform Prim Care*. 2014; 21: 161–170. <https://doi.org/10.14236/jhi.v21i4.72> PMID: 25479346
21. Ram P, Grol R, Rethans JJ, Schouten B, van der Vleuten C, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ*. 1999; 33: 447–454. <https://doi.org/10.1046/j.1365-2923.1999.00348.x> PMID: 10354322
22. Nilsen S, Baerheim A. Feedback on video recorded consultations in medical teaching: why students loathe and love it—a focus-group based qualitative study. *BMC Med Educ*. 2005; 5: 28. <https://doi.org/10.1186/1472-6920-5-28> PMID: 16029509
23. Rouquette A, Falissard B. Sample size requirements for the internal validation of psychiatric scales. *Int J Methods Psychiatr Res*. 2011; 20: 235–249. <https://doi.org/10.1002/mpr.352> PMID: 22020761
24. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 17: 101–110. [https://doi.org/10.1002/\(sici\)1097-0258\(19980115\)17:1<101::aid-sim727>3.0.co;2-e](https://doi.org/10.1002/(sici)1097-0258(19980115)17:1<101::aid-sim727>3.0.co;2-e) PMID: 9463853

25. Bland JM, Altman DG. Correlation, regression, and repeated data. *BMJ*. 1994; 308: 896. <https://doi.org/10.1136/bmj.308.6933.896> PMID: 8173371
26. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007; 60: 34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012> PMID: 17161752
27. Nunnally JC, Bernstein I. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill; 1994.
28. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990; 107: 238–246. <https://doi.org/10.1037/0033-2909.107.2.238> PMID: 2320703
29. Hancock GR, Mueller RO. *Structural equation modeling: A second course*. CT: Information Age. Greenwich; 2006.
30. Boussat B, François P, Gandon G, Gai J, Seigneurin A, Perneger T, et al. Inconsistencies Between Two Cross-Cultural Adaptations of the Hospital Survey on Patient Safety Culture Into French. *J Patient Saf*. 2017. <https://doi.org/10.1097/PTS.0000000000000443> PMID: 29140887
31. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011; 64: 96–106. <https://doi.org/10.1016/j.jclinepi.2010.03.002> PMID: 21130355
32. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979; 86: 420–428. <https://doi.org/10.1037//0033-2909.86.2.420> PMID: 18839484
33. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol*. 2012; 8: 23–34. <https://doi.org/10.20982/tqmp.08.1.p023> PMID: 22833776
34. McGraw KO, P S. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996; 1: 30–46. <https://doi.org/10.1037/1082-989X.1.1.30>
35. Gjersing L, Caplehorn JRM, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol*. 2010; 10: 13. <https://doi.org/10.1186/1471-2288-10-13> PMID: 20144247
36. Perneger TV, Leplège A, Etter JF. Cross-cultural adaptation of a psychometric instrument: two methods compared. *J Clin Epidemiol*. 1999; 52: 1037–1046. [https://doi.org/10.1016/s0895-4356\(99\)00088-8](https://doi.org/10.1016/s0895-4356(99)00088-8) PMID: 10526997
37. Chong L, Taylor S, Haywood M, Adelstein B-A, Shulruf B. Examiner seniority and experience are associated with bias when scoring communication but not examination skills in Objective Structured Clinical Examinations (OSCEs) in Australia. *J Educ Eval Health Prof*. 2018. <https://doi.org/10.3352/jeehp.2018.15.17> PMID: 30016854
38. Bokken L, Rethans J-J, Jöbsis Q, Duvivier R, Scherpbier A, van der Vleuten C. Instructiveness of real patients and simulated patients in undergraduate medical education: a randomized experiment. *Acad Med J Assoc Am Med Coll*. 2010; 85: 148–154. <https://doi.org/10.1097/ACM.0b013e3181c48130> PMID: 20042841
39. Bokken L, Rethans J-J, Scherpbier AJJA, van der Vleuten CPM. Strengths and weaknesses of simulated and real patients in the teaching of skills to medical students: a review. *Simul Healthc J Soc Simul Healthc*. 2008; 3: 161–169. <https://doi.org/10.1097/SIH.0b013e318182fc56> PMID: 19088660
40. Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *AAMC. Acad Med J Assoc Am Med Coll*. 1993; 68: 443–453. <https://doi.org/10.1097/00001888-199306000-00002> PMID: 8507309
41. Collins RMH J. P. AMEE Medical Education Guide No. 13: real patients, simulated patients and simulators in clinical examinations. *Med Teach*. 1998; 20: 508–521. <https://doi.org/10.1080/01421599880210>

CHAPITRE 2

Déterminants des compétences relationnelles en consultation médicale

3. CHAPITRE 2 : Déterminants des compétences relationnelles en consultation médicale

3.1. Partie 3) Compétences relationnelles en médecine générale ambulatoire

Article soumis — Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en médecine générale ambulatoire

Dans cet article soumis à la revue *Exercer*, nous avons cherché à valider les propriétés psychométriques de l'échelle 4-HCS en langue française dans le contexte de la médecine générale ambulatoire, puis à comprendre les spécificités des compétences relationnelles en soins primaires.

Objectifs

Les compétences relationnelles sont des compétences non techniques indispensables dans l'exercice des soins primaires. L'évaluation et la compréhension de ces compétences sont un enjeu pour l'amélioration de la qualité des soins. L'échelle SCQH est un outil standardisé permettant de mesurer ces compétences en médecine. L'objectif principal de cette étude était de valider cette échelle au sein de la médecine générale ambulatoire. Secondairement, l'étude cherchait à décrire ce niveau de compétence chez les médecins généralistes selon leur niveau d'expérience, leur sexe et les durées des consultations.

Méthodes

Il s'agissait d'une étude observationnelle transversale incluant des médecins généralistes en ambulatoire en Isère, Gard, Savoie et Haute-Savoie, ainsi que des patients consultant leur médecin traitant. Pour permettre l'évaluation des compétences relationnelles par l'échelle SCQH, les consultations étaient enregistrées puis analysées par 2 investigateurs. Un score total pour chaque item a pu être établi. La validité, la consistance interne et la fiabilité de l'échelle ont pu être étudiées.

Résultats

Le score total moyen des items était de 84,1/115 (écart type : 16,6). Le coefficient alpha de Cronbach représentant la consistance interne était excellent à 0,953 [IC95% : 0,93 ;0,96]. La

corrélation entre les items était en moyenne de 0,69 (écart-type : 0,13). La reproductibilité de la mesure était bonne avec un coefficient de corrélation intra-classe de 0,82 [IC95% : 0,74 ;0,88] et de 0,99 [IC95% : 0,97 ;0,99] représentant respectivement la fiabilité inter- et intra-juge. Le niveau de compétence augmentait avec la durée de la consultation ($p<0,001$), si le médecin était une femme ($p=0,002$) et pour les niveaux d'expérience plus faibles ($p=0,022$).

Conclusion

L'échelle SCQH est valide pour l'évaluation des compétences relationnelles en médecine générale ambulatoire et peut être transposable dans ce cadre à des fins d'évaluation et d'amélioration des compétences.

Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en médecine générale ambulatoire

Lara Pollier^{1*}, Romain Tailland^{1*}, Lucie Bosméan¹, Christel Oddou¹, Philippe Chaffanjon², Alexandre Bellier^{2,3,4}

¹ Univ. Grenoble Alpes, Département de Médecine Générale

² Univ. Grenoble Alpes, Faculté de Médecine

³ Service d'Épidémiologie et Évaluation médicale, CHU Grenoble Alpes

⁴ Équipe Biologie Computationnelle et Mathématique, Laboratoire TIMC, UMR 5525, CNRS, Univ. Grenoble Alpes

*Auteurs ayant contribué de manière égale à ce travail

Auteur correspondant :

Dr Alexandre Bellier

ORCID : 0000-0003-0907-0315

Service d'Épidémiologie et d'Évaluation Médicale

CHU Grenoble Alpes, CS10217

38043 Grenoble Cedex 09, France

E-mail : abellier@chu-grenoble.fr

Conflits d'intérêt : Les auteurs déclarent n'avoir aucun conflit d'intérêt.

Mots-clés : relations relationnelles, communication, médecine générale, psychométrie

INTRODUCTION

Les compétences relationnelles ont une importance majeure pour établir une relation de confiance entre le médecin et son patient avec un impact sur la qualité des soins, en améliorant la satisfaction des patients (1) et leurs observances aux traitements (2). Elles représentent d'ailleurs la compétence la plus désirée par les patients lors d'une consultation (3).

Il a été démontré qu'il était possible d'évaluer ces compétences non-techniques via l'utilisation d'instruments standardisés et validés, mais la plupart concernent l'évaluation d'une dimension spécifique,

comme la communication ou l'empathie (4). D'après Candib, les compétences relationnelles sont définies comme la présence de comportement verbaux et non-verbaux dans un contexte d'interactions individuelles avec le patient ou avec ses proches (5). Elles comprennent ainsi une dimension d'empathie et de communication, mais aussi de structuration de l'entretien qui est rarement pris en compte (6).

Le système de Codage des Quatre Habitudes (SCQH) est un instrument d'évaluation des compétences relationnelles qui couvre ces différentes dimensions des compétences relationnelles.

Elle a été développée par l'US Kaiser Permanente Organization, validée (7) puis largement utilisée à des fins d'enseignement aux États-Unis. Cette échelle d'évaluation a principalement été étudiée en secteur hospitalier. Une seule étude a utilisé cette échelle en médecine générale aux États-Unis (8). Par la suite, l'échelle SCQH a été traduite et validée dans plusieurs pays européens : en Norvège en 2010 (9), en Allemagne en 2014 (10) et en France récemment (6). La validation française a été effectuée, après adaptation transculturelle, dans le cadre de la formation médicale initiale des étudiants en médecine, sur la base de 200 enregistrements vidéo de simulations de consultations médicales (6).

Le SCQH est composé de 23 items évalués sur une échelle de Likert à 5 niveaux (Figure 1). Des scores plus élevés indiquaient de meilleures compétences relationnelles médecin-patient. Le niveau médian (c.-à-d. 3) et les niveaux finaux (c.-à-d. 1 et 5) bénéficient de descriptions comportementales spécifiques. Les évaluateurs sont encouragés à utiliser les autres niveaux (c.-à-d. 2 et 4) uniquement s'ils pensent que les compétences relationnelles se situent entre les niveaux précédents (Annexe 1). Les items sont divisés en 4 dimensions basées sur une manière organisée de penser et d'agir lors

d'une rencontre clinique centrée sur le patient : « S'investir dès le début », qui se concentre sur la création d'un lien de confiance rapide et la planification de la consultation ; « Obtenir le point de vue du patient », suscitant la compréhension du problème par le patient, les attentes du patient pour la visite, en déterminant l'impact de son problème sur sa vie quotidienne ; « Faire preuve d'empathie », qui encourage l'acceptation des émotions du patient, en l'aidant à identifier ses émotions et en utilisant un comportement non verbal approprié et « S'investir jusqu'à la fin », contenant des éléments axés sur la prise de décision et une information efficace (6,7).

L'échelle d'évaluation SCQH n'a pas encore été validée dans un contexte de médecine générale ambulatoire en France et rares sont les échelles d'évaluation à avoir exploré le contexte du soin primaires (4). Un tel changement de contexte d'utilisation nécessite une vérification des propriétés psychométrique de l'échelle, en termes de validité et de fiabilité (reproductibilité) pour pouvoir être utilisée en soin primaire. L'évaluation des compétences relationnelles peut non seulement aider le médecin à améliorer sa pratique en identifiant ses difficultés, mais elle peut également aider à comprendre les facteurs associés aux niveaux de compétences relationnelles.

Ainsi, l'objectif principal de l'étude était d'évaluer les propriétés psychométriques de l'échelle SCQH en soins primaires, en utilisant des données originales vidéo-enregistrées de consultation en médecine générale. Les objectifs secondaires visaient à décrire les niveaux de compétences relationnelles des médecins généralistes, à comprendre les déterminants parmi les caractéristiques des médecins ou de la consultation et à comparer les niveaux de compétences relationnelles par rapport à d'autre contexte de soins ou professions.

METHODES

Design de l'étude

Nous avons réalisé une étude observationnelle transversale à recrutement prospectif. L'étude a été réalisé en cabinet de médecine générale ambulatoire de juin 2020 à février 2021 sur 4 départements français (Haute Savoie, Savoie, Isère et Gard).

Critères d'éligibilité

L'étude incluait des médecins diplômés en Médecine Générale après information écrite et recherche de consentement. Les enregistrements de leur consultation concernaient des patients majeurs, parlant et comprenant le français, après signature d'un

consentement écrit. Les inclusions pouvaient concerner tout patient, quel que soit son motif de consultation.

Les critères d'exclusion concernaient les médecins non thésés, les refus a posteriori d'analyses des consultations et les problèmes techniques empêchant une interprétation complète et représentative sur la base de la consultation vidéo-enregistrée.

Recueil des données

Un enregistrement vidéo était réalisé via une caméra miniaturisée avec un objectif grand angle. L'intérêt de ce choix de caméra était d'analyser les réactions entre le patient et le clinicien, ainsi que d'obtenir une vision générale de l'environnement et de l'interaction avec celui-ci (11). L'hétéroévaluation réalisée sur vidéo permettait une évaluation complète des compétences relationnelles moins intrusive qu'avec un observateur présent durant la consultation et permettait de répéter l'évaluation (11,12).

Par voie d'annonces dans la subdivision de Grenoble, plusieurs médecins ont pu être recrutés. La sélection des journées de consultation a été réalisée au préalable en accord avec le médecin, puis le consentement de chaque patient était demandé au cabinet de médecine générale par les investigateurs. Le matériel d'enregistrement vidéo était installé préalablement à chaque consultation. Les

investigateurs n'étaient pas présents lors de l'entretien.

L'analyse des compétences relationnelles s'effectuait ensuite par deux investigateurs expérimentés, en double aveugle, sur la base de l'ensemble des enregistrements vidéo des consultations médicales. Un échantillon aléatoire de 30 consultations vidéo a également été sélectionné et analysé une deuxième fois par le même évaluateur pour évaluer la fiabilité intra-juge. En moyennant les évaluations de chaque investigateur, un score pour chaque item était établi puis la somme des scores aux items correspondait au score total, reflétant le niveau de compétence relationnelle.

Critères de jugement

Pour la validation de l'échelle SCQH en médecine générale, les propriétés psychométriques suivantes ont été étudiées : la fiabilité inter-juge, la fiabilité intra-juge et la cohérence interne avec calcul du coefficient alpha de Cronbach.

Le critère de jugement secondaire était le score total à l'échelle SCQH, reflétant le niveau de compétence relationnelle. Pour répondre aux objectifs secondaires, ce score a été évalué en fonction de la durée de la consultation, du statut de primo-consultant ou non pour chaque patient, de l'expérience du médecin générale (supérieure à 10 ans ou non) et du département d'exercice. Ces résultats ont été comparés aux données

d'évaluation issues de la même échelle pour d'autres publics (étudiants en médecine, sages-femmes, médecin hospitalier) en France ou à l'étranger.

Analyse statistique

En assumant un coefficient de corrélation intra-classe pour la mesure de la reproductibilité inter-opérateur et intra-opérateur de 0,70 avec 2 évaluateurs, nous estimons qu'un échantillon de 81 patients (consultations vidéo-enregistrées) serait suffisant pour démontrer avec une puissance de 80% un coefficient supérieur à 0,85 avec un seuil de significativité bilatéral à 5% (13).

Pour tenir compte de la variabilité inter-juge, les analyses des médecins sur la moyenne des scores pour chaque item entre les évaluateurs ont été réalisées (14). Nous avons procédé à des analyses descriptives pour le score total et pour chaque dimension du score issu de l'échelle SCQH. Les caractéristiques de l'échantillon d'étude ont également été décrits via un effectif et un pourcentage pour les variables qualitatives. Pour les variables quantitatives, une moyenne et un écart-type ont été utilisés après vérification de la normalité de la distribution des données ou à défaut une médiane et un écart interquartile.

La consistance interne a été évaluée par les mesures de corrélation inter-items et le coefficient alpha de Cronbach. C'est en

effet les paramètres les plus couramment utilisés pour valider la consistance interne d'une échelle d'évaluation (15).

Pour la fiabilité et donc la reproductibilité de cette échelle, le coefficient de corrélation intra-classe a été utilisé (16).

L'évaluation des objectifs secondaires a été réalisée par un modèle de régression linéaire pour variable dépendante quantitative continue, avec ajustement sur l'ensemble des caractéristiques dont l'association statistique p était inférieure à 0,20. Les analyses statistiques ont été réalisées via le logiciel RStudio (Version 1.0.14© 2009-2016 Rstudio, Inc.).

RESULTATS

Description de la population d'étude

Sur 88 consultations enregistrées sur une période de 6 mois, 80 ont été incluses dans l'étude, parmi 10 médecins différents, répartis dans 4 départements distincts. Huit consultations enregistrées ont été exclues de l'étude, dont deux a posteriori par les patients durant la semaine suivant la consultation et six pour des problèmes d'ordre technique audiovisuel. La population des médecins participants à l'étude était composée de 5 hommes et 5

femmes, âgés en moyenne de 43 ans (écart type : 14,4). Un médecin était considéré expérimenté s'il avait plus de 10 ans d'expérience professionnelle. Dans l'étude, ces derniers représentaient 60% des médecins (N=6). Les consultations enregistrées étaient réparties en 2 catégories selon qu'il s'agissait d'une première consultation ou d'une consultation avec un patient déjà connu par le médecin. 38% (N = 30) des consultations concernaient un premier contact avec le médecin consulté. Enfin, la durée moyenne d'une consultation était de 14,8 minutes (écart-type : 6.43) par consultation.

Validité et fiabilité de l'échelle

Sur l'échelle complète, le coefficient Alpha de Cronbach était de 0,953 (IC95% = 0,93-0,96) (Tableau 1). La moyenne des corrélations inter-items était de 0,69 (écart-type : 0,13).

Le Coefficient de corrélation intra-classe (ICC), reflétant la reproductibilité de la mesure, était de 0,82 (IC95% = 0,74-0,88) sur le score global pour la reproductibilité inter-évaluateur. Concernant la reproductibilité intra-évaluateur (test-retest), l'ICC était de 0,99 (IC95% = 0,97-0,99) sur l'échelle globale (Tableau 1).

Habitude – item	Moyenne (Ecart-type)		Corrélation inter-items moyenne	ICC inter-juge	ICC intra-juge	Alpha de Cronbach
1.S’investir dès le début	21,4	(5,10)	0,59	0,82	0,97	0,82
1- Montre de la familiarité	3,59	(0,97)	0,44			
2- Accueil le patient	3,79	(0,92)	0,41			
3- Bavardage	3,64	(1,35)	0,40			
4- Questions ouvertes	3,46	(1,14)	0,83			
5- Encourage l’expression des inquiétudes	3,63	(1,00)	0,74			
6- Suscite les préoccupations	3,27	(1,55)	0,74			
2.Obtenir le point de vue du patient	10,4	(2,59)	0,64	0,76	0,97	0,73
7- Compréhension du patient	3,33	(1,04)	0,77			
8- Demande les attentes	3,40	(1,07)	0,59			
9- Impact sur la vie quotidienne	3,69	(1,10)	0,57			
3. Faire preuve d’empathie	14,1	(3,33)	0,81	0,83	0,93	0,89
10- Encourage l’expression des émotions	3,71	(0,97)	0,86			
11- Accepte les sentiments	3,58	(0,88)	0,76			
12- Nomme les sentiments	2,74	(0,99)	0,77			
13- Comportement non verbal efficace	4,09	(0,97)	0,86			
4. S’investir jusqu’à la fin	38,2	(7,58)	0,71	0,78	0,97	0,92
14- Utilise le référentiel du patient	4,44	(0,84)	0,76			
15- Marque des pauses	4,58	(0,86)	0,78			
16- Explique sans jargon	4,56	(0,67)	0,71			
17- Explique les examens et les traitements	4,33	(1,02)	0,68			
18- Vérifie la compréhension	3,71	(0,98)	0,75			
19- Encourage la participation	3,64	(1,17)	0,78			
20- Vérifie l’acceptance	3,72	(1,12)	0,85			
21- Explore les obstacles	2,52	(1,04)	0,61			
22- Encourage les questions	3,27	(0,92)	0,66			
23- Etablit un plan de suivi	3,46	(1,22)	0,56			
Total	84,1	(16,6)	0,69			

Tableau 1. Propriétés métrologiques et psychométriques de l’échelle d’évaluation des compétences relationnelles
ICC=Coefficient de corrélation intra-classe (mesure de la reproductibilité)

Déterminants des compétences relationnelles en médecine générale

Le score moyen total chez les médecins généralistes est dans notre étude de 84,1/115 soit 14,6/20. Les items 13 « fait preuve d'un comportement verbal non efficace » à 17 « explique la nécessité des examens et traitements », obtiennent les meilleurs scores totaux sur l'ensemble des médecins, alors que ce sont les items 12 « aide à identifier et nommer les sentiments » et 21 « explore les obstacles à la mise en œuvre du plan thérapeutique », qui obtiennent les moins bons scores (Tableau 1).

Par ailleurs, les scores totaux à l'échelle SCQH sont augmentés chez les médecins avec la plus faible expérience clinique (80,8 versus 88,0). Cette différence était significative ($p=0,022$) après ajustement sur

le sexe, la durée de la consultation et le caractère primo-consultant ou non du patient (Tableau 2).

Les scores de compétences relationnelles variaient également selon le sexe du médecin avec des scores significativement plus élevés chez les femmes (91,6 vs 77,0, $p=0,002$).

La durée de consultation était quant à elle significativement associée au score à l'échelle SCQH ($p < 0,001$). Une durée de consultation plus longue était associée à un meilleur score à l'échelle SCQH avec un coefficient de corrélation de 0,498.

Enfin, le score SCQH était plus élevé en cas de consultation d'un patient primo-consultant qu'en cas de suivi, respectivement 88,4 vs 81,6, sans différence significative ($p= 0,624$).

		Score total	Univarié		Multivarié	
		SCQH	Coefficient β	p	Coefficient β	p
		Moyenne	(IC 95%)		(IC 95%)	
		(écart-type)				
Sexe	Femme (n=39)	91,6 (8,45)	-14,7 (21,3 ; -8,01)	< 0,001	-8,05 (-13,15 ; -2,94)	0,002
	Homme (n=41)	77,0 (19,2)				
Expérience clinique (> 10 ans)	Oui (n=43)	80,8 (20,2)	-7,13 (14,40 ; 0,14)	0,054	6,14 (0,89 ; 11,39)	0,022
	Non (n=37)	88,0 (9,99)				
Suivi du patient	Oui (n=50)	81,6 (19,5)	-6,81 (-14,30 ; 0,70)	0,075	1,60 (-4,87 ; 8,07)	0,624
	Non (n=30)	88,4 (8,54)				
Durée de la consultation	1 ^e quartile (3-10)	71,6 (18,5)	1,28 (0,78 ; 1,79)	< 0,001	0,63 (0,28 ; 0,98)	< 0,001
	2 ^e quartile (11-14)	88,4 (9,02)				
	3 ^e quartile (15-19)	90,8 (13,9)				
	4 ^e quartile (20-30)	91,3 (12,2)				

Tableau 2. Score de compétences relationnelles en fonction des caractéristiques de la consultation

DISCUSSION

Interprétation des résultats psychométriques

L'utilisation d'instruments standardisés validés est préconisée pour évaluer les compétences relationnelles des médecins (7). C'est dans ce cadre que s'inscrivait l'étude actuelle qui était la première à s'intéresser à la transposabilité de l'échelle SCQH en médecine générale. La cohérence interne de l'échelle dans ce contexte était satisfaisante avec des coefficients alpha de Cronbach systématiquement supérieur à 0,7 quel que soit la dimension. En comparaison avec les précédentes études françaises sur cette échelle, la cohérence interne était comparable (coefficient alpha de 0,94 et 0,97) (6,17), tout comme avec la seule étude américaine en soins primaire (coefficient alpha de 0,90) (8).

Par ailleurs, l'étude des corrélations inter-items montrait que seuls les items 4 (Utilise principalement des questions ouvertes), 10 (Encourage l'expression des émotions) et 13 (Fait preuve d'un comportement verbal non efficace) peuvent être considérés comme redondant entre eux, comme en atteste leur coefficient de corrélation supérieur à 0,8. L'item 20 (Vérifie l'acceptance du plan thérapeutique) pouvait aussi être redondant avec les items 18 (Vérifie efficacement la compréhension) et

19 (Encourage la participation à la prise de décision) dans ce contexte.

En ce qui concerne la fiabilité, cette dernière était excellente car le coefficient de corrélation intra classe est supérieur à 0,8 que ce soit en fiabilité inter et intra-juge (respectivement 0,82 et 0,99). L'échelle fournit donc des évaluations reproductibles, avec une meilleure reproductibilité inter-évaluateur que dans d'autres contextes d'utilisation où la reproductibilité était proche de 0,7 (6). Cependant, notre étude montre une meilleure reproductibilité intra-juge que dans des études précédentes (6,17), pouvant s'expliquer par une réévaluation avec un intervalle rapproché. De plus, l'évaluation de l'empathie (3^e dimension) semblait être la plus sensible au changement en test-retest, possiblement lié au faible nombre d'items dans cette dimension. En tenant compte de la précédente validation complète de l'échelle en langue française (6), nous pouvons ainsi conclure que l'échelle SCQH est valide et fiable pour l'évaluation des compétences relationnelles en médecine générale ambulatoire.

Interprétation des scores de compétences relationnelles en médecine générale

L'étude des scores de compétences relationnelles montrait une tendance pour les médecins généralistes à être plus à l'aise avec la dernière dimension de l'échelle SCQH, « S'investir jusqu'à la fin »,

évaluant la manière de conclure par l'annonce diagnostique, puis la prise en charge et la capacité du médecin à interroger le patient sur sa compréhension. Au contraire, la dimension « Obtenir le point de vue du patient » était la dimension qui mettait le médecin le plus en difficulté. Cette dernière concerne la capacité à s'intéresser à la compréhension du problème du patient, à lui demander ses attentes pour la consultation et à montrer de l'intérêt à propos de l'impact sur la vie quotidienne. Le même constat pouvait être fait dans la première étude de Krupat et al (7).

Par comparaison avec les précédentes études utilisant cette échelle (6,17), l'étude actuelle montre que les médecins généralistes sont parmi ceux avec les meilleurs résultats. L'étude actuelle obtenait un score total de 84,1 sur 115 (écart-type : 16,6), supérieur à la moyenne de l'ensemble des études qui est de 69,8. L'étude réalisée chez les sages-femmes (17) utilisant le SCQH retrouvait le meilleur score toutes études confondues, avec un score total de 98,7. Chez les étudiants en médecine (6), le score était de 76,4 (6). L'étude américaine impliquant les médecins généralistes (8) retrouvait un score très proche de 83,1 alors que les études impliquant des médecins hospitaliers avaient un score de 60,3 (9) à 68 (7). Ces résultats pourraient s'expliquer par la

proximité du médecin généraliste avec son patient, liée à une connaissance souvent ancienne du patient et de sa famille, par des consultations plus fréquentes, avec un temps dédié aux soins centré sur le patient plus important. En France l'organisation du soin est centrée sur le médecin traitant, qui est coordinateur de la santé de son patient (18). L'« approche centrée patient, relation, communication » est au centre des six compétences transversales recouvrant l'ensemble des tâches et fonctions attendues du médecin généraliste (19) et enseigné dans la formation universitaire des internes de médecine générale en France. Le médecin généraliste est habituellement le premier contact avec le système de soins, permettant un accès ouvert et non limité aux usagers, prenant en compte tous les problèmes de santé. Il développe une approche centrée sur la personne dans ses dimensions individuelle, familiale, et communautaire (20).

Interprétation des déterminants des compétences relationnelles

L'étude des objectifs secondaires montrait que les médecins les moins expérimentés bénéficiaient de scores totaux augmentés en moyenne de 8 points ($p=0,022$). Une précédente étude avait identifié cette tendance (21), qui peut s'expliquer par une durée de consultation plus longue pour ces derniers (22) et des formations facultaires

récentes autour d'une approche de soins centrée autour du patient. L'expérience clinique semble avoir un effet limité, raccourcissant souvent la durée de la consultation (23). Ces données suggèrent l'importance de la formation des internes en médecine générale et l'intérêt d'une formation médicale continue. Par ailleurs, les femmes avaient des niveaux de compétences relationnelles plus élevés dans notre étude, en accord avec la littérature. Il a été montré que les médecins de sexe féminin posaient plus de questions, font plus de retours a posteriori au patient, sourient et acquiescent plus en réponse au patient (24). Quant aux patients, ils parlent plus et révèlent plus d'information biomédicale et psychosociale aux médecins femmes (25).

Bien qu'il semble exister une association entre le score moyen du SCQH et le caractère de primo-consultant du patient, celle-ci n'était pas significative, sans doute liés en partie au fait que les primo-consultants étaient la plupart du temps retrouvés chez des médecins moins expérimentés, souvent remplaçants.

Limites de l'étude

Le médecin sachant qu'il était filmé pouvait avoir un comportement différent lors de la consultation et être influencé par la présence de la caméra (12). Toutefois, l'utilisation de caméra miniature avait

certainement limité cet effet. De plus, la majorité des médecins recrutés étaient impliqués dans la formation des jeunes médecins et étaient volontaires pour participer à l'étude, pouvant laisser penser que ces médecins étaient déjà sensibilisés à la qualité de la communication et du relationnel au sein de leur pratique.

L'étude présentait également des limites techniques avec l'absence d'enregistrement vidéo de certaines parties de la consultation qui pouvaient être parfois en dehors du champ de vision de la caméra. C'était le cas de l'accueil du patient notamment qui se déroule souvent en salle d'attente. Par ailleurs, en fonction de la typologie de la consultation, certains items de l'échelle étaient difficilement applicables. Ainsi, dans le cas d'un patient primo-consultant par exemple, les premiers items de l'échelle SCQH concernant la capacité du médecin à connaître son patient que ce soit sur le plan personnel ou médical n'était pas évaluable.

CONCLUSION

L'échelle SCQH d'évaluation des compétence relationnelles en consultation a été validée dans l'exercice de la médecine générale ambulatoire pour permettre une mesure reproductible. Les médecins généralistes avaient un niveau de compétence relationnelle se situant parmi

les meilleurs par rapport aux autres professionnels médicaux. Ces compétences relationnelles étaient d'autant plus importantes que le médecin était une femme, avec peu d'expérience clinique et réalisant des consultations plus longues que la moyenne. L'identification de ces facteurs associés aux compétences relationnelles doit nous questionner sur la formation des médecins généralistes en formation initiale, mais aussi en formation continue pour éviter le déclin de compétences essentielles à la qualité des soins. Dans ce contexte, l'échelle SCQH pourrait servir d'outil de

formation dans l'amélioration des compétences relationnelles en soins primaires. L'utilisation d'une évaluation standardisée de ces compétences sur la base de consultations enregistrées, qu'elles soient simulées ou effectuées avec des patients en cabinet de médecine générale, pourraient cibler les compétences sur lesquelles le médecin est en difficulté. L'évaluation de telles formation devraient faire l'objet d'étude de haut niveau de preuve pour attester de leur efficacité, dans l'intérêt des médecins et de leurs patients.

BIBLIOGRAPHIE

1. Boissy A, Windover AK, Bokar D, Karafa M, Neuendorf K, Frankel RM, et al. Communication Skills Training for Physicians Improves Patient Satisfaction. *J Gen Intern Med.* 2016 Jul;31(7):755–61.
2. Haskard Zolnierok KB, DiMatteo MR. Physician Communication and Patient Adherence to Treatment: A Meta-Analysis. *Med Care.* 2009 Aug;47(8):826–34.
3. Ong LML, de Haes JCJM, Hoos AM, Lammes FB. Doctor-patient communication: A review of the literature. *Soc Sci Med.* 1995 Apr;40(7):903–18.
4. Zill JM, Christalle E, Müller E, Härter M, Dirmaier J, Scholl I. Measurement of Physician-Patient Communication—A Systematic Review. Ozakinci G, editor. *PLoS ONE.* 2014 Dec 22;9(12):e112637.
5. Candib LM. When the doctor takes off her clothes: Reflections on being seen by patients in the exercise setting. *Fam Syst Health.* 1999;17(3):349–63.
6. Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J. Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills. Manzoni GM, editor. *PLOS ONE.* 2020 Apr 16;15(4):e0230672.

7. Krupat E, Frankel R, Stein T, Irish J. The Four Habits Coding Scheme: Validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns*. 2006 Jul;62(1):38–45.
8. Clayton MF, Latimer S, Dunn TW, Haas L. Assessing patient-centered communication in a family practice setting: How do we measure it, and whose opinion matters? *Patient Educ Couns*. 2011 Sep;84(3):294–302.
9. Fossli Jensen B, Gulbrandsen P, Dahl FA, Krupat E, Frankel RM, Finset A. Effectiveness of a short course in clinical communication skills for hospital doctors: Results of a crossover randomized controlled trial (ISRCTN22153332). *Patient Educ Couns*. 2011 Aug;84(2):163–9.
10. Scholl I, Nicolai J, Pahlke S, Kriston L, Krupat E, Härter M. The German version of the Four Habits Coding Scheme – Association between physicians' communication and shared decision making skills in the medical encounter. *Patient Educ Couns*. 2014 Feb;94(2):224–9.
11. Leong A, Koczan P, De Lusignan S, Sheeler I. A framework for comparing video methods used to assess the clinical consultation: a qualitative study. *Med Inform Internet Med*. 2006 Dec;31(4):255–65.
12. Coleman T. Using video-recorded consultations for research in primary care: advantages and limitations. *Fam Pract*. 2000 Oct;17(5):422–7.
13. Sd W, M E, A D. Sample size and optimal designs for reliability studies. *Stat Med* [Internet]. 1998 Jan 15 [cited 2021 Apr 14];17(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/9463853/>
14. Bland JM, Altman DG. Statistics Notes: Correlation, regression, and repeated data. *BMJ*. 1994 Apr 2;308(6933):896.
15. Thorndike RM. Book Review : Psychometric Theory (3rd ed.) by Jum Nunnally and Ira Bernstein New York: McGraw-Hill, 1994, xxiv + 752 pp. *Appl Psychol Meas*. 1995 Sep 1;19(3):303–5.
16. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011 Jan;64(1):96–106.
17. Blanc G, Bellier A, Di Marco L. Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en maïeutique. *Sages-Femmes*. 2020 Jul 1;19(4):50–4.

18. Médecin traitant et parcours de soins coordonnés [Internet]. [cited 2021 Apr 29]. Available from: <https://www.ameli.fr/assure/remboursements/etre-bien-rembourse/medecin-traitant-parcours-soins-coordonnes>
19. Compagnon L, Bail P, Huez J-F, Stalnikiewicz B, Ghasarossian C, Zerbib Y, et al. Définitions et descriptions des compétences en médecine générale. Vo U M E. :8.
20. Europe W, Wonca Europe. La définition européenne de la médecine générale - Médecine de famille. WONCA Europe 2002 (Les définitions européennes des caractéristiques de la discipline de médecine générale, du rôle du médecin généraliste et une description des compétences fondamentales du médecin généraliste - médecin de famille) [Internet]. World family doctors Caring for people Europe; 2002. 52 p. Available from: <http://www.woncaeurope.org>
21. Ruiz Moral R, Rodríguez Salvador JJ, Pérula De Torres L, Prados Castillejo JA. [Evolution of the communication profile of family medicine residents]. *Aten Primaria*. 2002 Feb 28;29(3):132–41.
22. Valverde Bolívar FJ, Pedregal González M, Pérez Fuentes MF, Alcalde Molina MD, Torío Durántez J, Delgado Rodríguez M. Communication skills of tutors and family medicine physician residents in Primary Care clinics. *Aten Primaria*. 2016 Dec;48(10):632–41.
23. Wouda JC, van de Wiel HBM. The communication competency of medical students, residents and consultants. *Patient Educ Couns*. 2012 Jan;86(1):57–62.
24. Jefferson L, Bloor K, Birks Y, Hewitt C, Bland M. Effect of physicians' gender on communication and consultation length: a systematic review and meta-analysis. *J Health Serv Res Policy*. 2013 Oct;18(4):242–8.
25. Hall JA, Roter DL. Do patients talk differently to male and female physicians? A meta-analytic review. *Patient Educ Couns*. 2002 Dec;48(3):217–24.

3.2. **Partie 4)** Compétences relationnelles en maïeutique

Article publié — Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en maïeutique

Dans cet article publié le 25 juillet 2020 dans la revue Sages-Femmes, nous avons cherché à valider les propriétés psychométriques de l'échelle 4-HCS en langue française dans le contexte des consultations de suivi de grossesse des sages-femmes, puis à comprendre les spécificités des compétences relationnelles en maïeutique.

Objectifs

Les compétences relationnelles sont des compétences non techniques indispensables dans l'exercice d'une sage-femme, mais aucun outil n'était disponible pour l'évaluation de ces compétences en maïeutique. Le Système de Codage des Quatre Habitudes (SCQH) est un outil standardisé permettant de mesurer ces compétences. L'objectif principal de cette étude était de valider cette échelle dans un contexte de consultation maïeutique en secteur hospitalier. Secondairement, l'étude cherchait à décrire le niveau de compétence des sages-femmes dans les différentes dimensions de l'échelle.

Méthodes

Il s'agissait d'une étude observationnelle transversale incluant des sages-femmes de trois centres hospitaliers français, ainsi que les patients venant consulter pour suivi de grossesse. Pour permettre l'évaluation des compétences relationnelles par l'échelle SCQH, les consultations étaient enregistrées puis analysées par deux investigateurs. Un score total pour chaque item a pu être établi. La validité et la fiabilité de l'échelle ont pu être étudiées.

Résultats

Le score total moyen des items était de 98,7/115 (écart type : 15,7). Le coefficient alpha de Cronbach représentant la consistance interne était excellent à 0,97 [IC95% : 0,96 ;0,98]. La reproductibilité de la mesure était satisfaisante avec un coefficient de corrélation intra-classe de 0,72 et de 0,96 représentant respectivement la fiabilité inter- et intra-juge. Les principaux points forts des sages-femmes sont la clarté des explications du plan de soins et la

communication verbale ou non-verbale. Leurs principales difficultés concernent la personnalisation de la consultation et l'approche centrée sur la patiente.

Conclusion

L'échelle SCQH est valide pour l'évaluation des compétences relationnelles en maïeutique et peut être transposable dans ce cadre à des fins d'évaluation et d'amélioration des compétences.

Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en maïeutique

Gwendoline
Blanc^{a,b,*}
Sage-femme

Alexandre Bellier^{c,d}
Médecin au Centre
d'investigation clinique
plurithématique du CHU
Grenoble-Alpes

Lionel Di Marco^{b,e}
Sage-femme

^aHôpital Necker, AP-HP,
149 rue de Sèvres,
75015 Paris, France

^bDépartement de maïeutique,
UFR de médecine,
Université Grenoble-Alpes,
Domaine de la Merci,
38700 La Tronche, France

^cService d'évaluation médicale,
CHU Grenoble-Alpes,
avenue Maquis-du-Grésivaudan,
38700 La Tronche, France

^dÉquipe biologie
computationnelle
et mathématiques, TIMC-
IMAG, UMR 5525, CNRS,
Université Grenoble-Alpes,
Domaine de la Merci,
38700 La Tronche, France

^eÉquipe ThEMAS,
TIMC-IMAG, UMR 5525,
CNRS, Université Grenoble-
Alpes, 621 avenue Centrale,
38400 Saint-Martin-d'Hères,
France

*Auteur correspondant.

Adresse e-mail :
blanc.gwendoline@gmail.com
(G. Blanc).

Une étude a été menée afin de valider le premier outil de mesure des compétences relationnelles des sages-femmes en consultation. La validité du système de codage des quatre habitudes montre des propriétés psychométriques et une reproductibilité de la mesure satisfaisantes. Les sages-femmes ont obtenu des scores élevés. Leurs principales difficultés concernent la personnalisation de la consultation et l'approche centrée sur la patiente. Cette échelle peut être utilisée à des fins d'évaluation ou de formation en maïeutique.

© 2020 Elsevier Masson SAS. Tous droits réservés

Mots clés - compétence professionnelle ; maïeutique ; pratique professionnelle ; psychométrie ; relation soignant-patient

Validation of the Four-Habits Coding Scheme for the evaluation of relational skills in midwifery. A study was conducted to validate the first tool for measuring the relational skills of midwives in consultation. The validity of the Four-Habits Coding Scheme shows satisfactory psychometric properties and measurement reproducibility. Midwives scored high. Their main difficulties concern the personalisation of the consultation and the patient-centred approach. This scale can be used for assessment or training in midwifery.

© 2020 Elsevier Masson SAS. All rights reserved

Keywords - caregiver-patient relationship; midwifery; professional competence; professional practice; psychometrics

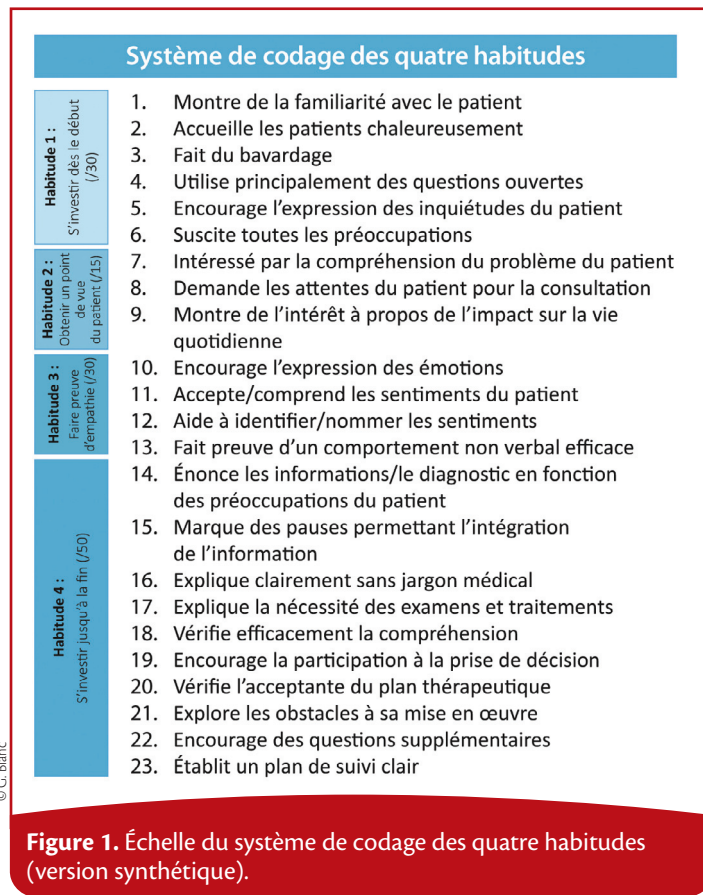
La qualité des relations entre les professionnels de santé et les patients est depuis longtemps décrite comme étant un déterminant majeur de la satisfaction du patient et de la réussite de sa prise en charge [1]. Il est en effet démontré que la qualité des interactions au cours de la consultation conditionne la satisfaction du patient, son adhésion sur le plan thérapeutique ou les résultats obtenus [1-3].

◆ À l'ère des nombreuses polémiques concernant les violences obstétricales, il semble très important de s'attacher à ce sujet [4]. Des compétences telles que la qualité de communication

et l'empathie sont un élément essentiel de la relation singulière qui se tisse entre la sage-femme et la patiente. En effet, les aptitudes relationnelles d'un praticien constituent un élément indispensable pour une médecine dont l'approche est centrée sur le patient [5,6]. Des études ont, par ailleurs, montré qu'il était possible d'améliorer ces qualités interpersonnelles par l'entraînement des professionnels de santé en cours de formation initiale ou continue [1,7].

◆ Dans ce contexte, il est nécessaire de disposer d'instruments standardisés, valides et fiables, pour évaluer, à des

fins formatives, ces compétences. La majorité de ces instruments ont été développés en langue anglaise et tous concernaient des médecins. Une échelle multidimensionnelle américaine a retenu l'attention de notre équipe du fait de ses propriétés psychométriques satisfaisantes. Développée par Edward Krupat *et al.*, elle est appelée 4-Habits Coding Scheme (4-HCS) ou, en français, système de codage des quatre habitudes (SCQH) (figure 1) [8]. Elle a récemment fait l'objet d'une adaptation transculturelle vers le français (annexe A) pour l'évaluation des qualités relationnelles des



médecins, mais n'a pas bénéficié d'une validation dans le domaine de la maïeutique [9].

◆ **Les qualités psychométriques d'un outil de mesure** sont évaluées grâce à la validité et à la fiabilité [10].

La fiabilité étudie la capacité de l'échelle à fournir des scores similaires, quels que soient les évaluateurs, les situations ou les périodes. Elle comprend la fiabilité interjuge (c'est-à-dire par des observateurs différents) et la fiabilité test-retest, ou fiabilité intrajuge (c'est-à-dire par le même observateur à des instants différents). Le coefficient utilisé dans ce contexte d'évaluation de la fiabilité d'une échelle est le coefficient de corrélation interclasse (ICC). Des résultats supérieurs à 0,70 sont considérés comme satisfaisants [11].

La validité s'intéresse notamment à la consistance interne de l'outil qui indique la capacité des items

à mesurer la même dimension ou le même construit. Elle est évaluée grâce au calcul du coefficient alpha de Cronbach. Plus il est élevé, plus on considère que les items sont homogènes entre eux. La consistance interne est satisfaisante si le coefficient alpha est supérieur ou égal à 0,70, il peut alors être considéré que les items du test sont similaires dans leur contenu (c'est-à-dire homogènes) [12-14].

◆ **L'objectif principal de notre étude¹ était donc de valider les propriétés psychométriques de l'échelle SCQH.** En effet, à ce jour, aucune échelle standardisée d'évaluation des compétences relationnelles des sages-femmes n'a été développée et validée en langue française.

Notre objectif secondaire était de décrire les compétences relationnelles des sages-femmes à l'aide de l'échelle, afin de mettre

en avant leurs points forts et leurs points faibles, et ce, à visée pédagogique.

Matériel et méthodes

Type d'étude

Cette étude de validation de l'échelle d'évaluation des qualités relationnelles SCQH est une étude observationnelle transversale, multicentrique, qui a été réalisée dans les services de consultations obstétricales de l'hôpital Couple-Enfant du centre hospitalier universitaire Grenoble-Alpes (38), du centre hospitalier de Valence (26) et du groupement hospitalier mutualiste de Grenoble (38). Elle s'est déroulée de septembre 2018 à avril 2019.

Recueil de données

◆ **Les données ont été recueillies prospectivement par observation** de consultations programmées de suivi de grossesse. Nous avons inclus entre trois et sept consultations par sage-femme.

◆ **Les consentements de la sage-femme et de la patiente** ont été obtenus de manière écrite, afin de réaliser un enregistrement audio anonymisé. Un avis éthique consultatif favorable a été obtenu le 19 septembre 2018².

◆ **Les observations ont été évaluées** une, deux ou trois fois. Elles l'ont été tout d'abord de façon directe, en consultation, par une étudiante sage-femme, qui a analysé les enregistrements une seconde fois, *a posteriori*. Ces évaluations ont été réalisées dans un intervalle de temps minimal de quinze jours. Un deuxième évaluateur a également écouté les enregistrements. Il s'agissait d'un médecin spécialiste des compétences relationnelles. Enfin, nous avons sélectionné de façon aléatoire dix consultations qui ont bénéficié d'une troisième évaluation.

Notes

¹ Cette étude a été réalisée par l'auteur dans le cadre de son mémoire pour l'obtention du diplôme d'État de sage-femme, en 2019.

² CECIC Rhône-Alpes-Auvergne, Clermont-Ferrand, IRB 5891.

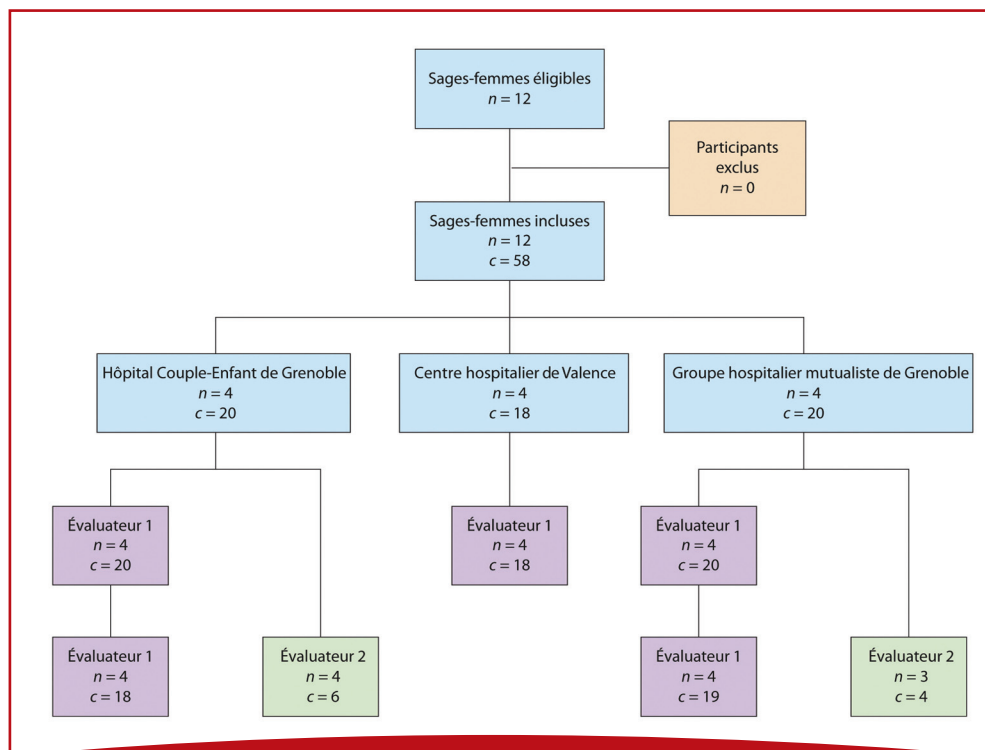


Figure 2. Diagramme de flux.
n : nombre de sages-femmes ; c : nombre de consultations.

52

© G. Blanc

Critères d'éligibilité

Les sages-femmes incluses étaient les professionnelles diplômées, effectuant des consultations programmées de suivi de grossesse auprès de patientes majeures, comprenant et parlant la langue française.

Critères de jugement et analyses

◆ **Le critère de jugement principal** était composé de :

- la cohérence interne évaluée par le coefficient alpha de Cronbach ;
- la fiabilité intrajuge évaluée en test-retest à l'aide du coefficient de corrélation intraclasse ;
- la fiabilité interjuge évaluée à l'aide du coefficient de corrélation intraclasse.

◆ **Le critère de jugement secondaire** était le score synthétique produit par l'échelle SCQH. Les analyses statistiques ont été effectuées sur le logiciel RStudio®.

Résultats

Description de l'échantillon

Au total, 12 sages-femmes ont été observées : quatre au centre hospitalier de Valence, quatre à l'hôpital Couple-Enfant de Grenoble et quatre au sein du groupe hospitalier mutualiste de Grenoble (figure 2). Finalement, avec les différentes consultations réalisées par chaque sage-femme et évaluées par les différents opérateurs, 105 évaluations ont été effectuées, dont 21, une seule fois.

Validité et fiabilité de l'échelle

◆ **L'évaluation de la cohérence interne de l'échelle** a été réalisée via le calcul du coefficient alpha de Cronbach. Tous les items avaient un coefficient alpha jugé satisfaisant, puisque supérieur à 0,70. Le coefficient alpha global pour l'échelle était de 0,97 (intervalle de confiance à 95 % : 0,96-0,98).

◆ **Concernant la fiabilité intrajuge** (test-retest), sur les 23 items évalués, nous avons retrouvé une fiabilité de 0,96, tandis qu'elle avait été estimée à 0,72.

Description des scores

◆ **L'objectif secondaire était de décrire les compétences relationnelles** des sages-femmes avec la nouvelle échelle, afin de mettre en avant leurs points forts et leurs points faibles.

◆ **Globalement, les scores obtenus sont bons**, puisque la moyenne globale est de 98,7/115 ($\pm 15,7$) (85,2 %), ce qui correspond à 17,2/20. Pour chaque dimension, les moyennes sont :

- habitude 1 - s'investir dès le début : 26,9/30, soit 17,9/20 (89,6 %) ;
- habitude 2 - obtenir le point de vue de la patiente : 11,9/15, soit 15,9/20 (79,33 %) ;
- habitude 3 - faire preuve d'empathie : 16,9/20 (84,5 %) ;
- habitude 4 - s'investir jusqu'à la fin : 43/50, soit 17,2/20 (86 %).

◆ **Les compétences relationnelles sont détaillées** dans le tableau 1. Pour chaque item, le professionnel a été évalué grâce à l'utilisation des catégories 1, 3 ou 5 (la catégorie 1 correspondant au score le plus faible et la catégorie 5 au score le plus élevé). Si l'évaluateur considérait que le praticien se situait entre deux catégories, il pouvait utiliser les notes 2 ou 4 (annexe A).

Les items n° 23 "établir un plan de suivi clair", n° 15 "marque des pauses permettant l'intégration de l'information", n° 16 "expliquer clairement sans jargon médical" et n° 13 "faire preuve d'un comportement non verbal efficace" sont ceux pour lesquels les scores obtenus sont les plus forts. Par ailleurs, l'item n° 21 "explorer les obstacles à la mise en place du plan thérapeutique" est

Tableau 1. Pourcentage des scores obtenus pour chaque item relatif aux compétences relationnelles des sages-femmes évalué.

Items évalués ^a	Catégories ^b				
	1	2	3	4	5
Item 1	0%	0%	2%	19%	79%
Item 2	0%	0%	10%	10%	79%
Item 3	0%	0%	12%	10%	78%
Item 4	0%	0%	17%	36%	47%
Item 5	0%	2%	21%	29%	48%
Item 6	0%	9%	9%	31%	52%
Item 7	2%	2%	21%	36%	40%
Item 8	2%	9%	19%	40%	31%
Item 9	3%	10%	22%	22%	41%
Item 10	0%	2%	19%	19%	60%
Item 11	0%	0%	26%	31%	43%
Item 12	5%	5%	29%	28%	33%
Item 13	0%	9%	7%	7%	78%
Item 14	0%	9%	7%	21%	64%
Item 15	0%	2%	10%	16%	72%
Item 16	0%	0%	2%	16%	83%
Item 17	0%	7%	16%	17%	60%
Item 18	0%	3%	12%	17%	67%
Item 19	0%	3%	14%	40%	43%
Item 20	0%	3%	17%	22%	57%
Item 21	16%	16%	41%	24%	3%
Item 22	2%	7%	22%	14%	55%
Item 23	0%	0%	0%	9%	91%

^aLa liste des items est disponible dans l'annexe A.

^bLa catégorie 1 correspond au score le plus faible et la catégorie 5 au score le plus élevé.

celui où les scores sont les plus faibles. En effet, 16 % des sages-femmes ont obtenu le score le plus bas avec "le clinicien ne mentionne pas les éventuels obstacles à la mise en œuvre du plan thérapeutique". Les autres items ayant les scores les moins élevés sont les items n° 8 "demande les attentes du patient pour la consultation", n° 9 "montre de l'intérêt à propos de l'impact sur la vie quotidienne" et n° 12 "aide à identifier ou nommer les sentiments" (tableau 1).

Discussion

Interprétation des propriétés psychométriques de l'échelle

◆ **Pour rappel, plus le coefficient alpha de Cronbach est élevé**, plus il est considéré que les items sont homogènes entre eux. Tous les items avaient un

coefficient alpha jugé satisfaisant puisque supérieur à 0,70 [12-14]. Le coefficient alpha global est même supérieur à 0,90, ce qui permet d'attester de la très bonne cohérence interne de notre échelle. Sa fiabilité a été évaluée grâce à la fiabilité interjuge et à la fiabilité intrajuge dont l'ICC était supérieur à 0,70. Ces bonnes propriétés lui permettent d'être utilisée pour l'évaluation des compétences relationnelles.

◆ **La validité externe des résultats est difficile à évaluer**, car les seules études visant à valider un outil d'évaluation des qualités relationnelles ont été réalisées à destination des médecins [11,15-18].

Interprétation des scores de compétences relationnelles obtenus

◆ **Il ressort de cette étude que les principaux points forts dont**

font preuve les sages-femmes sont les suivants : "établir un plan de suivi clair", "expliquer clairement sans jargon médical", "marquer des pauses permettant l'intégration de l'information" et "faire preuve d'un comportement non verbal efficace".

◆ **Le plan de suivi est régulier**, avec généralement une consultation par mois. L'orientation de la patiente, en cas de besoin, et l'organisation du suivi sont des compétences indispensables de la sage-femme. Des recommandations pour la pratique professionnelle ont d'ailleurs été éditées par la Haute Autorité de santé [19]. Il semble ainsi logique que cet item soit celui qui a obtenu les meilleures notes.

◆ **Concernant les trois autres items**, nous pouvons supposer, étant donné les scores obtenus, que les professionnels y sont naturellement sensibilisés,

Références

- [1] Haskard Zolnieriek KB, DiMatteo MR. Physician communication and patient adherence to treatment: A meta-analysis. *Med Care* 2009;47(8):826-34.
- [2] Garrity TF. Medical compliance and the clinician-patient relationship: A review. *Soc Sci Med* 1981;15(3):215-22.
- [3] Stewart MA. What is a successful doctor-patient interview? A study of interactions and outcomes. *Soc Sci Med* 1984;19(2):167-75.
- [4] Déchalotte M. *Le livre noir de la gynécologie*. Paris: Éditions First; 2017.
- [5] Ishikawa H, Hashimoto H, Kiuchi T. The evolving concept of "patient-centeredness" in patient-physician communication research. *Soc Sci Med* 2013;96:147-53.
- [6] Michie S, Miles J, Weinman J. Patient-centredness in chronic illness: what is it and does it matter? *Patient Educ Couns* 2003;51(3):197-206.
- [7] Van Zanten M, Boulet JR, McKinley D. Using standardized patients to assess the interpersonal skills of physicians: six years' experience with a high-stakes certification examination. *Health Commun* 2007;22(3):195-205.
- [8] Krupat E, Frankel R, Stein T, Irish J. The four habits coding scheme: Validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns* 2006;62(1):38-45.
- [9] Bellier A. *Adaptation transculturelle en langue française et validation des propriétés psychométriques de l'échelle 4-Habits Coding Scheme (4-HCS) pour l'évaluation des compétences relationnelles en consultation médicale* 2017.

Références

- [10] Mokkink LB, Terwee CB, Patrick DL, et al. The Cosmin checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual Life Res* 2010;19(4):539-49.
- [11] Marquer C. Validation transculturelle d'une échelle en contexte humanitaire. Dépistage des difficultés psychologiques du jeune enfant haoussa au Niger (Maradi) [Thèse de doctorat de psychologie]. Paris: Université René Descartes, Paris V; 2014. <https://tel.archives-ouvertes.fr/tel-01195996/document>.
- [12] Stewart M, Brown JB, Boon H, et al. Evidence on patient-doctor communication. *Cancer Prev Control* 1999;3(1):25-30.
- [13] Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the Cosmin checklist. *Qual Life Res* 2012;21(4):651-7.
- [14] Central Test. Validation Psychométrique. 2011. http://l.centraltest.fr/ct_fr/upload/documents/ValidationPsychometrie.pdf.
- [15] Bellier A, Chaffanjon P, Francois P, Labarère J. Content and psychometrics properties of standardized instruments measuring physician relational skills in consultation – A systematic review. 2017.
- [16] Zill JM, Christalle E, Müller E, et al. Measurement of physician-patient communication – A systematic review. *PLoS One* 2014;9(12):e112637.
- [17] Dictionnaire de français Larousse en ligne. Définition de psychométrie. www.larousse.fr/dictionnaires/francais/psychom%C3%A9trie/64855.
- [18] Caron J. Un guide de validation transculturelle des instruments de mesure en santé mentale. 1999. <http://instrumentspsychometriques.mcgill.ca/instruments/guide.htm>.
- [19] Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol* 2015;68(4):435-41.

Annexe A. Matériel complémentaire

Le matériel complémentaire (Annexe A) accompagnant la version en ligne de cet article est disponible sur www.sciencedirect.com et <http://dx.doi.org/10.1016/j.sagf.2020.06.013>.

Déclaration de liens d'intérêts
Les auteurs déclarent ne pas avoir de liens d'intérêts.

peut-être grâce à leur formation ou à leur vécu au cours de leur formation initiale ou du fait d'une expérience personnelle.

◆ **L'item "explorer les obstacles à la mise en place du plan thérapeutique"** est celui où les scores sont les plus mauvais. Il est possible que cela soit en lien avec le plan de suivi prédéfini. Ainsi, le professionnel, puisqu'il suit un déroulement classique de suivi sans mettre en place un plan thérapeutique spécifique (sauf exceptions), se questionne peu à propos des difficultés pouvant être rencontrées par la femme enceinte. De plus, dans des établissements de santé où les rendez-vous sont souvent peu nombreux, à des plages horaires très limitées, il peut être difficile de s'organiser en fonction de la patientèle. Ainsi, il semble légitime d'imaginer que les sages-femmes ne questionnent pas les patientes à ce sujet, puisqu'elles pourront difficilement adapter le plan thérapeutique. Il nous apparaît cependant utile de former les professionnels à l'importance de cet aspect, afin qu'ils puissent l'évoquer au cours des consultations. En effet, il semble indispensable de pointer les difficultés des femmes enceintes pour les aider à bénéficier d'un suivi optimal. Nous pouvons finalement penser qu'une meilleure prise en compte des obstacles rencontrés dans la mise en place du plan thérapeutique pourrait limiter le nombre d'absences aux consultations, et serait bénéfique tant aux professionnels qu'aux institutions. Ce résultat peut être mis en lien avec celui d'un autre item, qui

a également obtenu des scores relativement faibles : "montre de l'intérêt à propos de l'impact sur la vie quotidienne". Nous pouvons supposer que ces deux items vont plutôt de pair.

◆ **Les autres items ayant obtenu les scores les moins élevés** sont "demande les attentes du patient pour la consultation" et "aide à identifier ou nommer les sentiments". Le premier résultat peut certainement s'expliquer par le fait que les sages-femmes ont souvent un déroulé de consultation préétabli. Demander à la patiente si elle a des attentes particulières est susceptible d'obliger le professionnel à modifier son organisation. La solution pourrait donc être de poser la question à la patiente en fin de consultation, par exemple en lui demandant : « *Avez-vous d'autres attentes pour la consultation ?* »

◆ **Le deuxième résultat peut s'expliquer par le fait que certains professionnels préfèrent se limiter au domaine médical**, qu'ils maîtrisent parfaitement. Ils éprouvent peut-être la crainte de déclencher, chez la patiente, des émotions qu'ils n'auraient pas la capacité de gérer. En effet, la relation avec la femme enceinte peut se retrouver modifiée par le contenu de ce qu'elle révèle et, parfois, remettre en question l'objectivité du praticien, dans la mesure où cela le ramène à ses propres vécus affectifs. Il nous apparaît intéressant de disposer d'outils permettant de déterminer ces difficultés. Cela permet de pointer tant l'importance de la formation des sages-femmes, afin

qu'elles soient capables de réagir face aux émotions des patientes, que la nécessité de disposer d'un espace de parole pour mettre des mots sur les difficultés rencontrées par les professionnels dans la relation soignant-soigné.

Conclusion

Nous avons validé l'échelle SCQH afin de disposer du premier outil d'évaluation en langue française des compétences relationnelles des sages-femmes en consultation. Les propriétés psychométriques sont satisfaisantes, aussi bien en termes de validité que de reproductibilité. Cette échelle répond à la problématique d'une évaluation objective de données subjectives. Elle pourrait être, après identification des points forts et des points faibles des professionnels, une base pour la mise en place de formations théoriques et pratiques visant à améliorer les compétences de chacun, tant en formation initiale qu'à des fins de formation continue. Nous pourrions ainsi imaginer l'utiliser auprès des étudiants sages-femmes dans le cadre de consultations simulées, comme cela a été réalisé auprès d'étudiants en médecine [15]. Il serait aussi possible d'organiser des ateliers destinés aux sages-femmes en exercice, durant lesquels seraient diffusés des extraits filmés de consultations permettant aux professionnels, en évaluant les qualités de leur pair, de se questionner sur leur propre pratique. ●

3.3. **Partie 5)** Facteurs associés aux compétences relationnelles chez les étudiants en médecine en consultations simulées

Article publié — Student characteristics associated with interpersonal skills in medical consultations

Dans cet article publié le 3 mai 2022 dans la revue BMC Medical Education, nous avons cherché à explorer les facteurs associés aux compétences relationnelles. Le public concerné était des étudiants en médecine de quatrième année avec une évaluation des compétences relationnelles permis par l'échelle 4-HCS validée en langue française, dans le cadre de simulations avec patients standardisés.

Contexte

La qualité des soins médicaux dépend d'une communication efficace entre le médecin et le patient. Les compétences relationnelles peuvent être améliorées par l'enseignement, mais les déterminants du développement de ces compétences restent mal compris. Nous avons donc évalué les facteurs associés aux compétences relationnelles des étudiants en médecine lors de consultations médicales simulées.

Méthodes

Nous avons mené une étude transversale sur des étudiants en médecine de quatrième année participant à des consultations simulées avec des patients standardisés. Chaque consultation médicale vidéo-enregistrée a été évaluée indépendamment par deux évaluateurs, à l'aide d'une adaptation transculturelle en français de l'échelle 4-Habits Coding Scheme (4-HCS). Nous avons ensuite recueilli des informations sur les caractéristiques démographiques et sur les caractéristiques liées au parcours universitaire. La relation entre le score global du 4-HCS et les caractéristiques des étudiants a été modélisée à l'aide d'une régression linéaire univariée puis multivariée.

Résultats

Notre échantillon d'analyse comprenait 165 étudiants en médecine. Les facteurs significativement associés au score 4-HCS étaient le sexe ($\beta=-4,8$, $p=0,011$) et la réalisation d'un stage clinique à l'international ($\beta=6,2$, $p=0,002$) ou d'un stage en laboratoire de recherche ($\beta=6,5$, $p=0,005$). Les caractéristiques liées au parcours universitaire, notamment aux examens

par questions à choix multiples de la première à la troisième année de médecine, ainsi que le nombre de stages hospitaliers d'externat, en service de médecine ou de chirurgie, n'étaient pas significativement associés au score 4-HCS.

Conclusions

Les étudiants de premier cycle avec le meilleur niveau de compétences relationnelles lors de consultations médicales vidéo-enregistrées avec des patients standardisés étaient plutôt des femmes, ayant effectuées un stage clinique à l'international dans le cadre du programme d'échange ERASMUS ou un stage en laboratoire de recherche.

RESEARCH

Open Access



Student characteristics associated with interpersonal skills in medical consultations

Alexandre Bellier^{1,2,3,4*}, Philippe Chaffanjon¹, Patrice Morand¹, Olivier Palombi¹, Patrice Francois^{1,2} and José Labarère^{1,2,3}

Abstract

Background: The quality of medical care depends on effective physician–patient communication. Interpersonal skills can be improved through teaching, but the determinants are poorly understood. We therefore assessed the factors associated with the interpersonal skills of medical students during simulated medical consultations.

Methods: We conducted a cross-sectional study of fourth-year medical students participating in simulated consultations with standardized patients. Each video-recorded medical consultation was independently assessed by two raters, using a cross-cultural adaptation of the Four Habits Coding Scheme (4-HCS) into French. We then collected information on demographics and education-related characteristics. The relationship between the overall 4-HCS score and student characteristics was modeled using univariable and multivariable linear regression.

Results: Our analytical sample included 165 medical students for analysis. The factors significantly associated with 4-HCS score were gender ($\beta = -4.8, p = 0.011$) and completion of an international clinical placement ($\beta = 6.2, p = 0.002$) or a research laboratory clerkship ($\beta = 6.5, p = 0.005$). Education-related characteristics, multiple-choice examinations in the first to third preclinical years, and number of medicine or surgery clerkships were not significantly associated with 4-HCS score.

Conclusions: Undergraduate students with higher level of interpersonal skills during video-recorded medical consultations with standardized patients are more likely to be female, to have completed international clinical placement as part of the ERASMUS exchange program or research laboratory clerkship.

Keywords: Interpersonal skills, Doctor–patient relationship, Pedagogy, Simulation, Evaluation

Background

Effective physician–patient communication is essential for high-quality medical consultations [1]. Physician communication and interpersonal skills influence patient satisfaction [2], adherence to the care plan [3], and health outcomes [4–6]. Better communication and interpersonal skills would also reduce the likelihood of litigation

and malpractice claims [7] and have positive effects on physicians themselves, including greater job satisfaction and decreased risk of burnout [8, 9].

According to the Kalamazoo II report [10], communication and interpersonal skills are two distinct components of effective medical consultations. Communication skills refer to the performance of specific observable tasks, “that include interviewing to obtain a medical history, explaining a diagnosis and prognosis, giving therapeutic instructions and information needed for informed consent to undergo diagnostic

*Correspondence: abellier@chu-grenoble.fr

⁴ CIC 1406, INSERM, University Grenoble Alpes, Grenoble, France
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

and therapeutic procedures, and providing counseling to motivate participation in therapy or to relieve symptoms” [10]. Interpersonal skills are “inherently relational” and represent “a set of abilities required to successfully interact and communicate with others, both verbally and non-verbally”. Interpersonal skills encompass respectful attitude, attention paid to the patient, being personally present in the moment with the patient, interest in patient values and concerns, and real-time adjustment of the relationship [10].

Evidence is accumulating that communication and interpersonal skills can be acquired and improved through teaching and practice-based training for medical students [11–13]. Hence, communication and interpersonal skills training has been advocated as part of the educational curriculum for undergraduate medical students [12, 14–18]. Yet, current evidence on effective approaches to teaching communication skills remains limited [19].

Empirical research has focused on students’ attitude toward communication skills [20–23]. Indeed, attitude toward communication skills is assumed to predict to what extent students will invest in learning these skills and how they will use them when dealing with individual patients [24]. By contrast, few studies examined medical student demographics and education-related characteristics that were associated with observed communication and interpersonal skills. Identifying these factors may help faculty members understand the determinants of medical student communication skills and develop communication and interpersonal skills training programs that take student population characteristics into account [20]. They may include female gender [25–27], academic performance during preclinical years [28], early clinical experience [29–35], and participation in elective programs like international exchange or research laboratory placement.

To deal with this gap in knowledge, the aim of this study was to identify student demographics and education-related characteristics associated with communication and interpersonal skills assessed throughout medical consultations with standardized patients using the Four Habits Coding Scheme (4-HCS). Our primary research hypothesis was a gender difference in 4-HCS scores, consistent with literature. The secondary hypotheses were that 1) better academic performance during preclinical years reflected stronger student motivation and subsequent higher level of communication task fulfillment during medical consultations; 2) advancing clinical experience was associated with higher 4-HCS scores; and 3) participation in elective international exchange or research programs contributed to student interpersonal skills.

Methods

Study design

As part of a cross-sectional study [36], we analyzed the original data from video-recorded medical consultations with standardized patients at the University of Grenoble Alpes School of Medicine, France. The present manuscript complies with the STrengthening the Reporting of OBServational studies in Epidemiology (STROBE) guidelines [37].

Participants

Students who were registered in the fourth year of the medical curriculum were eligible to participate in the project if they passed the medical college admission test. Students were excluded from the present study if they were unavailable at the time of the simulated medical consultations because of leave for academic or personal reason, had been transferred from another school of medicine during their curriculum, or had applied without taking the medical college admission test.

Participants from two consecutive academic years were recruited from October 2016 to December 2016 and from February to April 2018, respectively. Participants were recruited on a voluntary basis.

Consultations with standardized patients

Each student completed a single video-recorded medical consultation with a standardized patient. For this purpose, seven medical consultation case scenarios were developed, based on the basic medical interview tasks that comprised the 4-HCS. For each medical consultation case scenario, seven clinical problems and seven profiles of portrayals were created according to personality type or character trait (Appendix 1).

The standardized patients consisted of 20 actors recruited at the Department of Performing Arts in Grenoble Alpes University. To ensure consistency in standardized patients, the 20 actors attended a two-hour session with provision of oral and written detailed instructions by a clinician for each medical consultation case scenario. The actors were then trained by their improvisation instructor to portray one to two cases and the profiles of portrayals were chosen from the list just before the consultation. All consultations were video-recorded and videos were stored in a secure cloud-database for subsequent remote access.

4-HCS scoring

Each video-recorded medical consultation was independently assessed by two raters, using the cross-cultural adaptation of the 4-HCS into French [36]. The 4-HCS is a standardized instrument designed to assess

23 basic medical interview tasks that are organized within four dimensions, including *Invest in the beginning* (six items), *Elicit the patient's perspective* (three items), *Demonstrate empathy* (four items), and *Invest in the end* (ten items) [38]. The 4-HCS overall score is computed by summing ratings for the individual items, ranging from 23 (i.e., less effective) to 115 (i.e., more effective). The French version of the 4-HCS previously demonstrated satisfactory internal consistency and acceptable reliability with two independent raters [36]. Cronbach's alpha was 0.94 for the 4-HCS overall score and intra-class correlation coefficient estimates based on average ratings were 0.82 and 0.91 for inter- and intra-rater reliability, respectively [36].

The two raters were a full professor of medicine and a resident in medicine, both with experience in simulation and teaching communication and interpersonal skills. They performed the assessments at home, usually in sessions of up to 2 consecutive hours. Data were captured using an online form in which the completion of each item of the 4-HCS was mandatory, so there could be no missing data. Because individual ratings were deemed unreliable [36], each student was assigned the average 4-HCS score from two independent raters.

Data collection

Information on demographics and education-related characteristics was electronically retrieved from databases maintained by the University of Grenoble Alpes School of Medicine. This included age, gender, number of medical college admission test attempts, the total score achieved in the basic science multiple-choice examinations for each of the first- to third-year academic class (range, 0–100), the number of internal medicine, surgery, pediatrics, psychiatry, as well as obstetrics and gynecology clerkship rotations completed, international clinical placement as part of the ERASMUS (EuRopean community Action Scheme for the Mobility of University Students) exchange program [39], and research laboratory placement. For students who took any examination more than once, only the score obtained at the first attempt was analyzed. ERASMUS and research laboratory placements were both elective. ERASMUS placement was conditional to foreign language level certification and the selection process was based on the grade point average achieved in second and third preclinical years. The pre-requisite for research laboratory placement was the validation of two elective disciplinary research courses and formal training on how conducting a research project. Students did not undergo Objective Structured Clinical Examinations (OSCEs) during the first three preclinical years.

Statistical methods

Summary statistics for the 4-HCS overall score were reported as means along with standard deviations (SDs). To quantify the univariable associations between medical student characteristics and 4-HCS score, crude (unadjusted) β regression coefficient point estimates along with 95% confidence intervals were derived from straight-line linear regression analysis for continuous dependent variable. β regression coefficients quantified the mean change in 4-HCS score for each independent variable category relative to the reference, with null value equal to 0. Adjusted β regression coefficients were derived from multiple linear regression analysis, using the same approach. β regression coefficients were adjusted for international clinical placement, research laboratory placement, and gender. All first-order interactions involving independent variables included in multiple linear regression model were tested for statistical significance.

Because medical consultation case scenarios were randomly assigned, we did not expect heterogeneity in 4-HCS overall scores according to medical problems and profiles of portrayals (Table S1). In unplanned exploratory analysis, we examined the univariable associations for each 4-HCS subscale score with baseline student characteristics.

Two-sided p -values of <0.05 were considered statistically significant. All analyses were performed using RStudio (Version 1.0.143©).

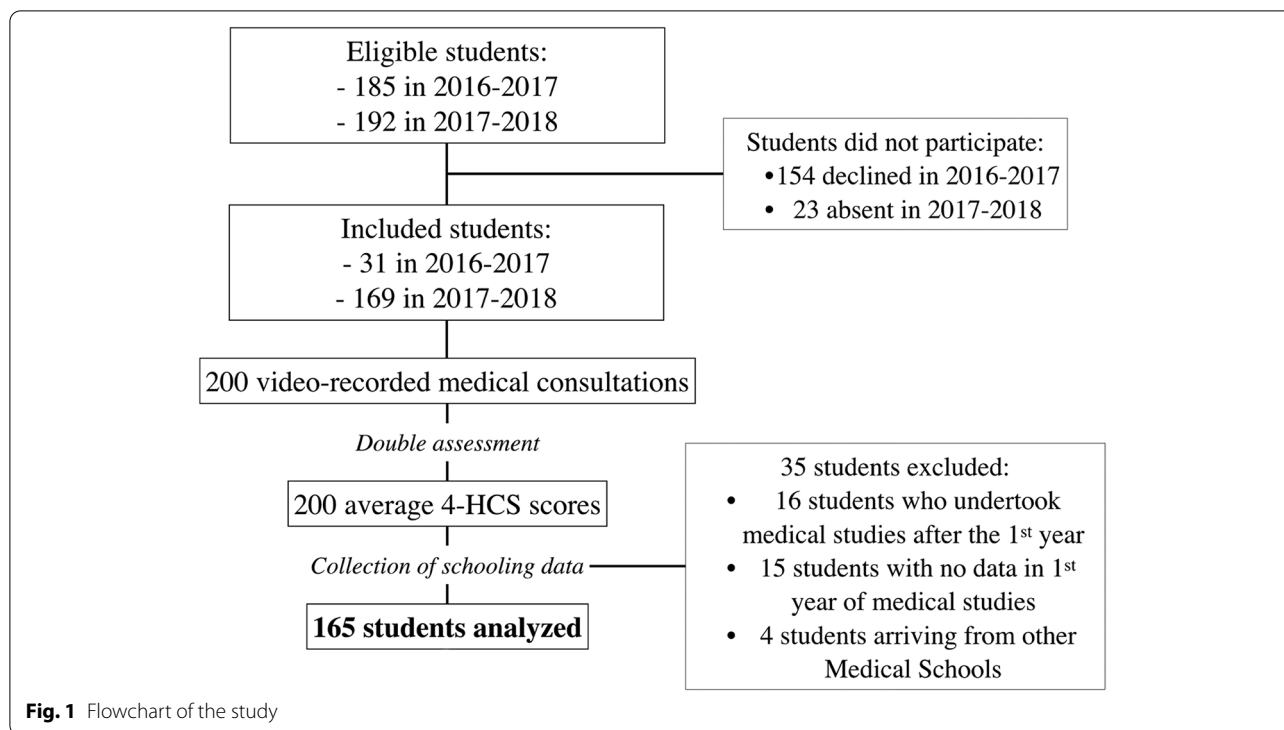
Ethical considerations

After appropriate written information was delivered, each student signed an individual consent form regarding participation in the research and image rights. An institutional review board (*Comité d'Ethique du Centre d'Investigation Clinique de Clermont-Ferrand*, IRB 00005891) reviewed and approved the study protocol and the information form prior to study initiation. All data were de-identified before analysis.

Results

Study sample

A total of 200 fourth-year medical students (31 and 169 in the 2016–2017 and 2017–2018 academic year, respectively) completed video-recorded consultations with standardized patients. After omitting 35 students according to the exclusion criteria (Fig. 1), our analytic sample consisted of 165 video-recorded medical consultations with standardized patients. The vast majority of participants were female (67% [111/165]) and the mean age was 21.8 years ($SD = 1.4$). Most students were admitted to the school of medicine after a second attempt (54% [89/165]) and the median



examination grades ranged from 73.6 in the first year to 66.8 in the third year. The median numbers of medicine and surgery clerkships were 3 (range, 0–5) and 1 (range, 0–4), respectively. Overall, 47 students (28.5%) had completed a research internship and 29 had participated in the ERASMUS international exchange program (17.6%). Female students had better examination grades, fewer research laboratory placement and comparable ERASMUS and hospital medical placement rates than their male counterparts (Table S2).

4-HCS scoring

The mean 4-HCS overall score was 75.4 (SD=11.8) for all participants. Intra-class correlation coefficient estimate for inter-rater reliability based on individual ratings was 0.76. In univariable analysis, female gender, completion of a research laboratory placement, and international clinical placement as part of the ERASMUS exchange program were associated with a higher 4-HCS overall score (Table 1). These associations remained significant in multivariable analysis (Table 2). Our study failed to show any significant relationship between 4-HCS overall score and the number of medical college admission test attempts, the total score achieved in basic science multiple-choice examinations, and the number of internal medicine and surgery clerkship rotations completed in both univariable and

multivariable analyses (Tables 1 and 2). No significant first-order interaction involving independent variables included in multiple linear regression was found.

No evidence of inconsistencies in univariable associations of baseline student characteristics was observed across 4-HCS subscale scores (Table S3), although no formal conclusions could be drawn from these exploratory analyses.

Discussion

This study provides important insight into undergraduate student characteristics that are potential determinants of communication performance observed during video-recorded medical consultations with standardized patients. The characteristics independently associated with 4-HCS overall score included female gender, completion of a research laboratory placement, and international clinical placement. The number of medical college admission test attempts, basic science examination scores, and the number of medicine or surgery clerkship rotations completed were not predictive of 4-HCS overall scores.

To our knowledge, limited data exist on the relationship between examination scores or number of medical college admission test attempts and communication skills for medical students. A previous study reported a weak but significant correlation between academic examination scores undertaken over the first 2 years

Table 1 Univariable associations of 4-HCS overall score with demographic and education-related characteristics for undergraduate medical students ($n = 165$)

Characteristics		4-HCS ^b Mean (SD)	β regression coefficient (95%CI) ^a		p
			Unadjusted		
Gender	Men ($n = 54$)	72.4 (13.5)	−4.4 (− 8.2;0.6)	Reference	0.024
	Women ($n = 111$)	76.8 (10.6)			
Research laboratory clerkship	Yes ($n = 29$)	79.6 (8.7)	5.1 (0.4;9.9)	Reference	0.032
	No ($n = 136$)	74.5 (12.2)			
International clinical placement	Yes ($n = 47$)	79.4 (10.4)	5.6 (1.7;9.5)	Reference	0.006
	No ($n = 118$)	73.8 (11.9)			
No. of attempts at MCAT	1 ($n = 79$)	76.7 (11.2)	2.5 (− 1.1;6.1)	Reference	0.180
	> 1 ($n = 86$)	74.2 (12.2)			
First year examination score	1st quartile [68.3–71.2]	73.0 (12.7)	5.6 (0.4;10.7)	Reference	0.866
	2nd quartile [71.3–73.5]	78.5 (11.5)			
	3rd quartile [73.6–75.8]	76.6 (11.1)			
	4th quartile [75.9–87.6]	74.2 (11.1)			
Second year examination score	1st quartile [51.8–66.4]	72.9 (11.4)	3.6 (− 1.5;8.7)	Reference	0.292
	2nd quartile [66.5–70.9]	76.5 (11.7)			
	3rd quartile [71.0–75.3]	77.0 (10.0)			
	4th quartile [75.4–87.0]	74.9 (10.6)			
Third year examination score	1st quartile [38.0–62.0]	73.3 (15.2)	2.0 (− 3.2;7.3)	Reference	0.302
	2nd quartile [62.1–66.7]	75.7 (10.8)			
	3rd quartile [66.8–71.7]	77.1 (9.9)			
	4th quartile [71.8–84.3]	75.3 (10.7)			
Medicine clerkship	No ($n = 2$)	85.1 (2.3)	−8.3 (−26.5;10.0)	Reference	0.038
	One ($n = 8$)	76.9 (10.3)			
	Two ($n = 35$)	77.6 (11.5)			
	Three ($n = 75$)	75.0 (11.8)			
	Four ($n = 40$)	74.8 (11.6)			
Surgery clerkship	Five ($n = 5$)	64.5 (15.0)	−20.7 (−40.0;−1.4)	Reference	0.478
	No ($n = 28$)	74.1 (13.5)			
	One ($n = 80$)	77.0 (9.8)			
	Two ($n = 46$)	74.0 (13.1)			
	Three ($n = 10$)	71.5 (14.6)			
	Four ($n = 1$)	80.8	6.6 (−17.1;30.3)		

Abbreviations: MCAT Medical College Admission Test, CI Confidence Interval, 4-HCS 4 Habits Coding Scheme, SD Standard Deviation

^a β regression coefficients quantified the change in 4-HCS score for each independent variable category relative to the reference, with null value equal to 0

^b 4-HCS range: [23–115]

of medical school and subsequent OSCEs [28], which constitute an alternative approach for assessing student communication and interpersonal skills [40]. Our study failed to show any significant association between basic science multiple-choice examination scores and subsequent 4-HCS overall score. Although we cannot exclude that our study was underpowered to detect a significant correlation, this negative finding likely reflects the fact that communication and interpersonal skills are independent concepts from technical skills evaluated throughout standardized examinations. This latter

observation supports the incorporation of communication skills training and assessment as part of the medical curriculum beside standardized examinations on basic science or clinical knowledge.

International clinical placement as part of the ERASMUS exchange program was associated with higher 4-HCS overall scores in our study. Previous studies have suggested that international clinical placement would provide students with the opportunity to develop communication skills [41]. Yet, the potential for self-selection bias cannot be excluded, reflecting the fact that students

Table 2 Multivariable associations of 4-HCS overall score with demographic and education-related characteristics for undergraduate medical students ($n = 165$)

Characteristics		4-HCS ^b Mean (SD)	β regression coefficient (95%CI) ^a	
			Adjusted ^c	p
Gender	Men ($n = 54$)	72.4 (13.5)	-4.8 (-8.4;-1.1)	0.011
	Women ($n = 111$)	76.8 (10.6)	Reference	
Research laboratory clerkship	Yes ($n = 29$)	79.6 (8.7)	6.5 (1.9;11.0)	0.005
	No ($n = 136$)	74.5 (12.2)	Reference	
International clinical placement	Yes ($n = 47$)	79.4 (10.4)	6.2 (2.3;10.0)	0.002
	No ($n = 118$)	73.8 (11.9)	Reference	
No. of attempts at MCAT	1 ($n = 79$)	76.7 (11.2)	1.3 (-2.3;4.9)	0.480
	> 1 ($n = 86$)	74.2 (12.2)	Reference	
First year examination score	1st quartile [68.3–71.2]	73.0 (12.7)	Reference	0.717
	2nd quartile [71.3–73.5]	78.5 (11.5)	2.5 (1.1;10.9)	
	3rd quartile [73.6–75.8]	76.6 (11.1)	2.4 (-1.1;8.4)	
	4th quartile [75.9–87.6]	74.2 (11.1)	2.4 (-4.5;5.0)	
Second year examination score	1st quartile [51.8–66.4]	72.9 (11.4)	Reference	0.851
	2nd quartile [66.5–70.9]	76.5 (11.7)	3.9 (-0.9;8.7)	
	3rd quartile [71.0–75.3]	77.0 (10.0)	1.5 (-3.5;6.5)	
	4th quartile [75.4–87.0]	74.9 (10.6)	2.6 (-4.7;5.4)	
Third year examination score	1st quartile [38.0–62.0]	73.3 (15.2)	Reference	0.916
	2nd quartile [62.1–66.7]	75.7 (10.8)	1.5 (-3.3;6.4)	
	3rd quartile [66.8–71.7]	77.1 (9.9)	0.3 (-4.9;5.5)	
	4th quartile [71.8–84.3]	75.3 (10.7)	0.2 (-4.9;5.3)	
Medicine clerkship	No ($n = 2$)	85.1 (2.3)	Reference	0.256
	One ($n = 8$)	76.9 (10.3)	-12.0 (-29.7;5.7)	
	Two ($n = 35$)	77.6 (11.5)	-8.4 (-24.6;7.8)	
	Three ($n = 75$)	75.0 (11.8)	-10.5 (-26.6;5.5)	
	Four ($n = 40$)	74.8 (11.6)	-10.0 (-26.5;6.4)	
Surgery clerkship	Five ($n = 5$)	64.5 (15.0)	-18.1 (-36.9;0.8)	0.403
	No ($n = 28$)	74.1 (13.5)	Reference	
	One ($n = 80$)	77.0 (9.8)	4.1 (-0.8;9.1)	
	Two ($n = 46$)	74.0 (13.1)	3.0 (-2.5;8.6)	
	Three ($n = 10$)	71.5 (14.6)	3.1 (-5.5;11.6)	
	Four ($n = 1$)	80.8	9.9 (-12.7;32.5)	

Abbreviations: MCAT Medical College Admission Test, CI Confidence Interval, 4-HCS 4 Habits Coding Scheme, SD Standard Deviation

^a β regression coefficients quantified the change in 4-HCS score for each independent variable category relative to the reference, with null value equal to 0

^b 4-HCS range: [23–115]

^c Adjusted for international clinical placement, research laboratory clerkship and gender

with a higher level of communication skills at baseline were selected for international clinical placement and subsequently yielded better communication performance during medical consultations with standardized patients.

Research laboratory placement was a local initiative at our medical school and therefore the report of similar findings in the literature is unlikely. We hypothesize that the benefits of research laboratory placement include the development of teamwork and interpersonal skills. As part of their training on how conducting a research

project, students are strongly recommended to collaborate with multidisciplinary teams including laboratory technicians, clinical research associates and supervisors. Moreover, the final evaluation of research laboratory clerkship includes among others assessment of collaborative relationships by the project supervisor, a fact that may foster student interpersonal skills.

Surprisingly, we did not observe any significant relationship between the number of internal medicine or surgery clerkship rotations completed and 4-HCS scores.

Basically, we expected that interpersonal skills improved with hospital clinical experience. The scant results from the literature suggest an increase in interpersonal skills with an early immersion in an internship [29]. Yet a decline of empathy during the medical curriculum has been reported by several studies [31–35]. Although speculative, a potential explanation might be that the repetition of tasks, the work overload, and the high emotional load of certain clinical situations lead students pay less attention to interpersonal skills. Hence, advancing clinical experience may result in better accomplishment of communication tasks that is attenuated by decrease in interpersonal skills.

Our finding that female gender was associated with higher 4-HCS overall score is consistent with previous research indicating that female students tend to achieve higher grades in clinical communication tasks than their male counterparts [42].

Krupat et al. found that female medical students were more patient-centered [27]. This difference was observed very early in the medical curriculum and guided the career choice with a greater interest in primary care. These differences can be explained by women's increased interest in communication and listening in comparison with men [26], particularly in nonverbal communication [43]. Meta-analyses have revealed that consultation time is significantly greater for women than for men and that female physicians showed greater engagement in communication [44, 45]. Beyond communication, it has been shown that female medical students developed more empathy [32]. The choice of different career orientations is reflected in our results, since female medical students tend to have more hospital medical placements and fewer surgery placements. Female students also have significantly better academic results in the majority of disciplines. However, the academic results are not associated with the interpersonal skills measured by the 4-HCS scale. The discrepancy in the numbers of male versus female participants in our study reflected the larger proportion of females among medical students at our university. Reassuringly, the participation rate did not differ between male and female medical students (55% versus 71%, $P=0.481$).

This study has potential implications for optimizing communication skills training and assessment as part of the medical curriculum. First, objective measures such as multiple-choice examinations do not predict medical student communication performance, and the assessment of communication and interpersonal skills requires specific – although subjective – evaluations, such as consultations with standardized patients. Second, internal medicine clerkship should include other

training components than clinical reasoning and technical skills so as to provide students with the opportunity to acquire experience in communication with real patients. Third, male medical students may benefit from specific training, so as to gain the same level of communication skills as their female counterparts. Fourth, future research should examine the role of other important variables, including personality and interpersonal style, in the development of communication and interpersonal skills during medical studies.

The limitations of our study should be acknowledged. First, each student had the opportunity to conduct only one video-recorded consultation, and this may not reflect his / her communication performance in routine practice. Second, we may have omitted some relevant characteristics relating to 4-HCS scoring, including ethnic minority, socioeconomic status, or cognitive factors such as emotional intelligence. Importantly, attitude toward communication skills was not measured in our study, a potential determinant of observed communication. Third, our study was conducted at a single study site and our findings may not apply to other settings or geographical areas. Fourth, between-actor heterogeneity could not be investigated in this study, and we could not exclude that this heterogeneity might partly explain our findings despite random assignment. Fifth, inter-rater reliability (0.76) was lower than previously reported (0.82). This apparent inconsistency likely reflects different computational approaches, although between-study heterogeneity in rater level of experience and random sampling error cannot be excluded. Indeed, inter-rater reliability was quantified using average ratings over two out of four raters in the previous study [36]. Because only two raters participated in the present study, average ratings could not be computed. Reassuringly, intra-class correlation coefficient based on individual ratings in the present study was higher than 0.70, which is indicative of satisfactory inter-rater reliability.

Conclusions

Through medical consultation simulation exercises and standardized assessment of medical students' interpersonal skills, we showed that the main factors associated with interpersonal skills were gender and participation in an international clinical placement or a research laboratory clerkship without influence of education-related characteristics or hospital experience. Based on our results, we showed the importance of these simulation exercises with the assessment of nontechnical skills. Indeed, multiple-choice examinations do not predict the level of interpersonal skills, and hospital medicine clerkship does not seem to

improve them. It would be interesting to study the decline of empathy, especially among medical students. A major challenge for initial training would be to avoid this decline. It is likely that simulation exercises such as those we have organized will help students realize the importance of these skills and limit the unavoidable decline.

Abbreviations

4-HCS: Four Habits Coding Scheme; SD: Standard Deviation.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12909-022-03412-9>.

Additional file 1: Table S1. Medical consultation case scenarios and profiles of portrayals.

Additional file 2: Table S2. Comparison of undergraduate medical student characteristics according to gender ($n = 165$).

Additional file 3: Table S3. Exploratory analysis of univariable associations between 4-HCS subscale scores and demographic and education-related characteristics for undergraduate medical students ($n = 165$).

Acknowledgments

We thank Pauline Bouchet and Ariane Martinez, two professors of performing arts, for their collaboration in setting up simulated medical consultations with standardized patients. We thank Linda Northrup for critical reading and substantial editing of the manuscript.

Authors' contributions

AB, PC and JL conducted the study design and study supervision. AB and PC performed data acquisition. Analysis, interpretation of data, statistical analysis and drafting the manuscript were carried out by AB, PF and JL. Critical revision of the manuscript for important intellectual content was performed by PM and OP. All authors read and approved the final manuscript.

Funding

The statistical analysis was developed within the Grenoble Alpes Data Institute, which is supported by the French National Research Agency under the "Investissements d'avenir" program (ANR-15-IDEX-02).

Availability of data and materials

In accordance with the French guidelines, the data set will be archived for at least 10 years after publication. The data set will be available on Harvard Dataverse public repository: <https://doi.org/10.7910/DVN/RHZVKY>.

Declarations

Ethics approval and consent to participate

All procedures performed in this study involving human participants were in accordance with the ethical standards of France and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Enrolled patients provided written informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹University Grenoble Alpes, School of Medicine, Grenoble, France. ²Clinical Epidemiology Unit, Grenoble Alpes University Hospital, CS10217, 38043 Grenoble

Cedex 09, France. ³Computational and Mathematical Biology Team, TIMC-IMAG UMR 5525, CNRS, University Grenoble Alpes, Grenoble, France. ⁴CIC 1406, INSERM, University Grenoble Alpes, Grenoble, France.

Received: 14 June 2021 Accepted: 11 April 2022

Published online: 03 May 2022

References

- Ong LML, de Haes JCJM, Hoos AM, Lammes FB. Doctor-patient communication: a review of the literature. *Soc Sci Med*. 1995;40:903–18.
- Boissy A, Windover AK, Bokar D, Karafa M, Neuendorf K, Frankel RM, et al. Communication skills training for physicians improves patient satisfaction. *J Gen Intern Med*. 2016;31:755–61.
- Zolnieriek KBH, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care*. 2009;47:826–34.
- Kelley JM, Kraft-Todd G, Schapira L, Kossowsky J, Riess H. The influence of the patient-clinician relationship on healthcare outcomes: a systematic review and meta-analysis of randomized controlled trials. *PLoS One*. 2014;9:e94207.
- Stewart J. Northern practices offer a challenge but few MDs willing to make the move. *CMAJ*. 1995;153:667–8.
- Dwamena F, Holmes-Rovner M, Gaudin CM, Jorgenson S, Sadigh G, Sikorskii A, et al. Interventions for providers to promote a patient-centred approach in clinical consultations. *Cochrane Database Syst Rev*. 2012;12:CD003267.
- Brennan MD, Monson V. Professionalism: good for patients and health care organizations. *Mayo Clin Proc*. 2014;89:644–52.
- Löffler-Stastka H, Seitz T, Billeth S, Pastner B, Preusche I, Seidman C. Significance of gender in the attitude towards doctor-patient communication in medical students and physicians. *Wien Klin Wochenschr*. 2016;128:663–8.
- Engerer C, Berberat PO, Dinkel A, Rudolph B, Sattel H, Wuensch A. Specific feedback makes medical students better communicators. *BMC Med Educ*. 2019;19:51.
- Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, et al. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. *Acad Med*. 2004;79:495–507.
- Yedidia MJ, Gillespie CC, Kachur E, Schwartz MD, Ockene J, Chepaitis AE, et al. Effect of communications training on medical student performance. *JAMA*. 2003;290:1157–65.
- Aspegren K. BEME Guide No. 2: teaching and learning communication skills in medicine—a review with quality grading of articles. *Med Teach*. 1999;21:563–70.
- Smith S, Hanson JL, Tewksbury LR, Christy C, Talib NJ, Harris MA, et al. Teaching patient communication skills to medical students: a review of randomized controlled trials. *Eval Health Prof*. 2007;30:3–21.
- von Fragstein M, Silverman J, Cushing A, Quilligan S, Salisbury H, Wiskin C, et al. UK consensus statement on the content of communication curricula in undergraduate medical education. *Med Educ*. 2008;42:1100–7.
- Bachmann C, Abramovitch H, Barbu CG, Cavaco AM, Elorza RD, Haak R, et al. A European consensus on learning objectives for a core communication curriculum in health care professions. *Patient Educ Couns*. 2013;93:18–26.
- Brouwers M, Custers J, Bazelmans E, van Weel C, Laan R, van Weel-Baumgarten E. Assessment of medical students' integrated clinical communication skills: development of a tailor-made assessment tool. *BMC Med Educ*. 2019;19:118.
- Hargie O, Boohan M, McCoy M, Murphy P. Current trends in communication skills training in UK schools of medicine. *Med Teach*. 2010;32:385–91.
- Junod Perron N, Klöckner Cronauer C, Hautz SC, Schnabel KP, Breckwoldt J, Monti M, et al. How do Swiss medical schools prepare their students to become good communicators in their future professional careers: a questionnaire and interview study involving medical graduates, teachers and curriculum coordinators. *BMC Med Educ*. 2018;18:285.
- Koponen J, Pyörälä E, Isotalus P. Comparing three experiential learning methods and their effect on medical students' attitudes to learning communication skills. *Med Teach*. 2012;34:e198–207.
- Rees C, Sheard C. The relationship between medical students' attitudes towards communication skills learning and their demographic and education-related characteristics. *Med Educ*. 2002;36:1017–27.

21. Anvik T, Gude T, Grimstad H, Baerheim A, Fasmer OB, Hjortdahl P, et al. Assessing medical students' attitudes towards learning communication skills--which components of attitudes do we measure? *BMC Med Educ.* 2007;7:4.
22. Gude T, Tyssen R, Anvik T, Grimstad H, Holen A, Baerheim A, et al. Have medical students' attitudes towards clinical communication skills changed over a 12-year period? A comparative long-term study. *BMC Med Educ.* 2020;20:11.
23. Cleland J, Foster K, Moffat M. Undergraduate students' attitudes to communication skills learning differ depending on year of study and gender. *Med Teach.* 2005;27:246–51.
24. Anvik T, Grimstad H, Baerheim A, Bernt Fasmer O, Gude T, Hjortdahl P, et al. Medical students' cognitive and affective attitudes towards learning and using communication skills--a nationwide cross-sectional study. *Med Teach.* 2008;30:272–9.
25. Crandall SJ, Volk RJ, Loemker V. Medical students' attitudes toward providing care for the underserved. Are we training socially responsible physicians? *JAMA.* 1993;269:2519–23.
26. Fürstenberg S, Schick K, Deppermann J, Prediger S, Berberat PO, Kadmon M, et al. Competencies for first year residents - physicians' views from medical schools with different undergraduate curricula. *BMC Med Educ.* 2017;17:154.
27. Krupat E, Hiam CM, Fleming MZ, Freeman P. Patient-centeredness and its correlates among first year medical students. *Int J Psychiatry Med.* 1999;29:347–56.
28. Adam J, Bore M, McKendree J, Munro D, Powis D. Can personal qualities of medical students predict in-course examination success and professional behaviour? An exploratory prospective cohort study. *BMC Med Educ.* 2012;12:69.
29. Helmich E, Bolhuis S, Laan R, Koopmans R. Early clinical experience: do students learn what we expect? *Med Educ.* 2011;45:731–40.
30. Murden RA, Way DP, Hudson A, Westman JA. Professionalism deficiencies in a first-quarter doctor-patient relationship course predict poor clinical performance in medical school. *Acad Med.* 2004;79(10 Suppl):S46–8.
31. Neumann M, Edelhäuser F, Tauschel D, Fischer MR, Wirtz M, Woopen C, et al. Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Acad Med.* 2011;86:996–1009.
32. Tavakol S, Dennick R, Tavakol M. Empathy in UK medical students: differences by gender, medical year and specialty interest. *Educ Prim Care.* 2011;22:297–303.
33. Hojat M, Vergare MJ, Maxwell K, Brainard G, Herrine SK, Isenberg GA, et al. The devil is in the third year: a longitudinal study of erosion of empathy in medical school. *Acad Med.* 2009;84:1182–91.
34. Calzadilla-Núñez A, Díaz-Narváez VP, Dávila-Pontón Y, Aguilera-Muñoz J, Fortich-Mesa N, Aparicio-Marengo D, et al. Erosion of empathy during medical training by gender. A cross-sectional study. *Arch Argent Pediatr.* 2017;115:556–61.
35. Chen DCR, Kirshenbaum DS, Yan J, Kirshenbaum E, Aeltline RH. Characterizing changes in student empathy throughout medical school. *Med Teach.* 2012;34:305–11.
36. Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J. Cross-cultural adaptation of the 4-habits coding scheme into French to assess physician communication skills. *PLoS One.* 2020;15:e0230672.
37. von EE, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ.* 2007;335:806–8.
38. Krupat E, Frankel R, Stein T, Irish J. The four habits coding scheme: validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns.* 2006;62:38–45.
39. Parey M, Waldinger F. Studying abroad and the effect on international labour market mobility: evidence from the introduction of ERASMUS*. *Econ J.* 2011;121:194–222.
40. Roshal JA, Chefitz D, Terregino CA, Petrova A. Comparison of self and simulated patient assessments of first-year medical students' interpersonal and communication skills (ICS) during objective structured clinical examinations (OSCE). *BMC Med Educ.* 2021;21:107.
41. Imafuku R, Saiki T, Hayakawa K, Sakashita K, Suzuki Y. Rewarding journeys: exploring medical students' learning experiences in international electives. *Med Educ Online.* 2021;26:1913784.
42. Joekes K, Noble LM, Kubacki AM, Potts HWW, Lloyd M. Does the inclusion of "professional development" teaching improve medical students' communication skills? *BMC Med Educ.* 2011;11:41.
43. Vogel D, Meyer M, Harendza S. Verbal and non-verbal communication skills including empathy during history taking of undergraduate medical students. *BMC Med Educ.* 2018;18:157.
44. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: a meta-analytic review. *JAMA.* 2002;288:756–64.
45. Jefferson L, Bloor K, Birks Y, Hewitt C, Bland M. Effect of physicians' gender on communication and consultation length: a systematic review and meta-analysis. *J Health Serv Res Policy.* 2013;18:242–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



3.4. **Partie 6)** Impact de la durée de consultation sur les compétences relationnelles des médecins

Impact of consultation length on interpersonal skills between medical practitioner and patient: a cross-sectional study

Dans ce travail préliminaire, nous avons cherché à identifier une association entre la durée de la consultation et les compétences relationnelles médecin-patient. En effet, la durée de consultation est souvent contrainte et l'enjeu est de comprendre si une consultation plus longue ou plus courte pourrait modifier les compétences relationnelles médecin-patient. Sur la base des nombreuses consultations évaluées par l'échelle 4-HCS, nous avons cherché à déterminer si une consultation dans un temps limité était compatible avec un bon niveau de compétences relationnelles.

Contexte

L'idée que la durée de la consultation améliore directement le résultat de la consultation, notamment sur la qualité de la communication, reste incertaine. Nous avons donc évalué l'impact de la durée de la consultation sur les compétences relationnelles parmi des médecins généralistes, des médecins hospitaliers, des étudiants en médecine et des sages-femmes.

Méthodes

Sur la base d'études transversales et d'un essai contrôlé randomisé, nous avons analysé des consultations vidéo-enregistrées avec des patients réels (pour les généralistes, les médecins hospitaliers et les sages-femmes) ou des patients standardisés (pour les étudiants en médecine). Chaque consultation médicale a été évaluée indépendamment par deux médecins, en utilisant l'adaptation transculturelle en français du Four Habits Coding Scheme (4-HCS). Le motif de la consultation a été recueilli, en utilisant la Classification internationale des soins primaires - 3e révision (ICPC-3). La relation entre le score global 4-HCS et la durée de la consultation a été modélisée à l'aide d'une régression linéaire multivariée et d'une méthode utilisant des polynômes fractionnels.

Résultats

Notre échantillon comprenait 894 consultations pour 679 médecins praticiens analysés. La durée médiane des consultations était de 23,0 min pour les médecins et de 9,2 min pour les étudiants en médecine. Cette durée était significativement associée ($p < 0,01$) à un meilleur score 4-HCS dans toutes les dimensions de l'échelle. Pour les étudiants en médecine et jusqu'à 6,85 minutes de consultation, l'augmentation de 4-HCS était de 7,68 par minute, puis de 1,24 par minute. Pour les praticiens seniors, cette augmentation était de 3,15 par minute jusqu'à 13,40 minutes de consultation, puis de 0,14 par minute.

Conclusions

Nous avons montré que la durée de la consultation était associée au niveau de compétences relationnelles du médecin. En effet, chez les étudiants en médecine et les praticiens, la durée de la consultation augmentait significativement avec le score 4-HCS, surtout dans les premières minutes.

Impact of consultation length on interpersonal skills between medical practitioner and patient: an observational study

Alexandre Bellier^{1,2,3,4*}, MD, MSc; Joey Fournier⁴, MD; Alexane Olagnon², MSc; Lucie Bosméan¹, MD; Philippe Chaffanjon¹, MD, MSc; José Labarère^{1,2,3}, MD, PhD

¹ Univ. Grenoble Alpes, School of Medicine, France

² Clinical Epidemiology Unit, Grenoble Alpes University Hospital, France

³ Computational and Mathematical Biology Team, TIMC-IMAG UMR 5525, CNRS, Univ. Grenoble Alpes, France

⁴ CIC 1406, INSERM, Univ. Grenoble Alpes, France

*Corresponding author:

Alexandre Bellier, MD, MSc

Clinical Epidemiology Unit, Grenoble University Hospital, CS10217, 38043 Grenoble Cedex 09, France

Phone: +33 (0)4 76 76 87 67

E-Mail: abellier@chu-grenoble.fr

Short title: Consultation length on interpersonal skills

ABSTRACT

Background — The idea that consultation length directly improves consultation outcome, particularly on quality communication, remains unclear. We therefore assessed the impact of consultation length on interpersonal skills among general practitioners, hospital physicians, medical students and midwives.

Methods — Based on cross-sectional studies and a randomized controlled trial, we analyzed video-recorded consultations with real (for general practitioners, hospital physicians, and midwives) or standardized patients (for medical students). Each video-recorded medical consultation was independently evaluated by two physicians, using the cross-cultural adaptation of the Four Habits Coding Scheme (4-HCS) in French. The gender of the practitioner and the consultation length were assessed concurrently based on the consultation records. The videos were reviewed by two physicians to assess reason for encounter, using the International Classification of Primary Care - 3rd Revision (ICPC-3). The relationship between the overall 4-HCS score and consultation length was modeled using multivariable linear regression and fractional polynomial method.

Results — Our analytical sample included 894 consultations from 679 medical practitioners for analysis. The median consultation length was 23.0 min for physicians and 9.2 min for medical students. Consultation length was also significantly associated ($p < 0.01$) with better 4-HCS in all dimensions. For medical students and up to 6.85 minutes of consultation, the 4-HCS increase was 7.68 per minute, then 1.24 per minute. For medical practitioners, this increase was 3.15 per minutes up to 13.40 minutes of consultation, then 0.14 per minute.

Conclusions — We have shown the association between the length of the consultation and the level of interpersonal skills of the physician. Indeed, among medical students and medical practitioners, the duration of the consultation significantly increased the score of the 4-HCS, especially in the first minutes.

1. Introduction

Physician interpersonal skills are key components of effective medical consultations (1) and comprise core physician competences that are most desired by patients (2). Evidence has accumulated supporting the conclusion that high-quality communication relates with enhanced patient satisfaction (3), greater adherence to treatment (4), better health outcomes (5), increase quality of life (6) and decreased risk of malpractice claims (7).

Consultation time is an important resource in health care, and concerns about the impact of its length on quality of care (8,9) may apply to the quality of communication. A longer consultation improves process and diagnosis quality, empowers the patients, and increase lifestyle advice and preventive activities, while reducing the amount of prescribing and general practitioners stress (10–12). A longer consultation also encourages the expression and research of the patient's quality of life in relation to their pathology (13).

However, the idea that consultation length directly improves consultation outcome, particularly on quality communication, remains unclear. A cross-sectional study suggests an association between consultation time and patient-centeredness or joint decision making, with overall better physician-patient communication (14). On

the contrary, some studies suggest that there is no direct link between the length of consultation and communication, suggesting that the improvement in outcomes is due more to the extra time spent with the patient than to the quality of the communication (15). Another study suggests that there is no demonstrated link between the length of consultation and the patient's experience of communication, with a very positive feeling for some short consultations but overall a better quality of care for longer consultations (16). However studies on the subject remain of low level of evidence (17).

In addition, the duration of consultation is very heterogeneous between countries, from 48 seconds in Bangladesh to 22,5 minutes in Sweden. In France, the average consultation length is 16 minutes (18). This discrepancy has many explanations, the lack of doctors forces to increase the number of consultations per day (more than 90 in China) (19), physicians' factors such as gender and experience (20,21) and patients' factors like gender, diagnosis and social class (21–23) are likely to explain such differences. The environment (urban, rural and suburban) also has an impact on consultation length (24), as does the gender of the physician. Female physicians appear to have longer consultations and engage in more constructive behaviors such as

reassurance, positive talk and empathy. They also have more patient-centered talk (25–27). Finally, the reason for consultation leads to important disparities concerning the duration of consultation.

The medical profession is therefore faced with the need to have good communication skills over a defined consultation length. We hypothesize that consultation length is positively associated with higher interpersonal skills.

In the present study, we aimed to assess the consultation length impact on interpersonal skills among general practitioners, hospital physicians, medical students, and midwives. Secondary objectives were to assess the impact of gender, status of medical professional and reason for the encounter, on consultation length.

2. Methods

Study design

In cross-sectional studies (28,29) and a randomized controlled trial (30), we analyzed original data from consultations at the University of Grenoble Alpes Medical School and several hospitals or doctor's office in the Grenoble area (France). The present manuscript follows the Strengthening the Reporting of Observational studies in Epidemiology (STROBE)(31) guidelines.

Study sample

The study sample consisted of video-recorded medical consultations with real or standardized patients, including four types of health professionals: general practitioners (GP), hospital physicians (HP), midwives and medical students (22–24)

GP with an outpatient practices in four French geographic area (Gard, Haute-Savoie, Isère and Savoie) were included from June 2020 to February 2021. These physicians were asked to include about ten patients, regardless of reasons for encounter. The recordings concerned adult patients, speaking and understanding French, consulting for the first time or in follow-up.

HP practicing at the Grenoble Alps University Hospital were included from July to October 2021. Each physician volunteering to participate was board-certified in medical, surgical or gynecology-obstetrics specialty. Physicians were required to include about 8 initial consultations for new adult patients (30). No restrictions on French comprehension were applied. Otherwise, inclusion and exclusion criteria were the same as before.

Midwives practicing at the Grenoble Alps University Hospital, Valence hospital, or

the Mutualist Clinic in Grenoble, France, were included from September 2018 to April 2019 (29). Midwives were asked to include several pregnancy follow-up consultations. Restrictions on French comprehension were applied. Consultations were recorded after signing a written consent. Exclusion criteria were retrospective refusals and technical problems.

Students enrolled in the fourth year of the medical curriculum at the Grenoble Alpes University were included from 2017 to 2020. Students in 2017 were recruited on a voluntary basis. In subsequent years, each student recorded a video and it was part of their academic program. Students were instructed not to exceed ten minutes of consultation, although there was no time control. Patients were performed by other students, in University Grenoble Alpes Performing Arts Department, who had seven clinical scenarios previously established by the investigators. Exclusion criteria were for students who were on an exchange program during consultations session.

All medical consultations with standardized patients were video-recorded using professional video equipment. Video-recording of consultations overcame some of the challenges of direct observation (32)

Outcomes

Each video-recorded medical consultation was independently evaluated by two physicians, using the cross-cultural adaptation of the Four Habits Coding Scheme (4-HCS) in French (28).

The two raters were randomly assigned: a full professor of medicine, an assistant professor of medicine, a general practitioner, or a medical resident, all of whom had experience with simulation, teaching communication and interpersonal skills.

The gender of the practitioner and the consultation length were assessed concurrently based on the consultation records. The HP videos were reviewed by two physicians to assess reason for encounter, using the International Classification of Primary Care - 3rd Revision (ICPC-3)(33). These were the only videos not deleted after the first viewing.

4HCS description

We selected the Four Habits Coding Scheme (4-HCS) to assess interpersonal skills. The 4-HCS is a standardized instrument designed to assess 23 basic medical interviewing tasks that are organized within four dimensions, including *Invest in the beginning* (six items), *Elicit the patient's perspective* (three items), *Demonstrate empathy* (four

items), and *Invest in the end* (ten items). The 4-HCS overall score is computed by adding the scores for each item, ranging from 23 (i.e., less effective) to 115 (i.e., more effective) (34). The 4-HCS is based on the conceptual framework of the “Four Habits Model,” a training program that was developed within the US Kaiser Permanente Health Maintenance Organization and implemented to teach effective communication skills to thousands of clinicians in this organization over the two last decades (35).

The French version demonstrates satisfactory internal consistency with a Cronbach’s alpha of 0.94 and average intra-class correlation coefficient estimates were 0.82 and 0.91 for inter- and intra-rater reliability, respectively (28).

ICPC description

The World Organization of Family Doctors (WONCA) created the International Classification of Primary Care (ICPC) in 1987, and the third edition (ICPC-3) was published in 2020. The ICPC is widely used in general practice for the classification of three important elements of the health care encounter: reason for encounter, diagnoses or problems, and processes of care. ICPC is formally recognized by the World Health Organization’s Family of International Classifications as a classification for primary care. We chose the ICPC to assess

reason for encounter because, although it is designed for primary care, it appears to be the best classification for classifying the reason for encounter among several specialties.

It is divided into 19 chapters. There are four chapters for: reasons for visit, social problems, interventions and functioning related issues. Fourteen chapters represent the localization of the problem on a body-system level, each subdivided into two components (complaints and diagnoses). Another chapter includes the same components for problems that cannot be classified into one or more than two body systems.

In this study, we grouped the ICPC codes on body system level subdivided into complaints and diagnoses. A second classification divided the ICPC codes in three reasons for encounter: consultation for a symptom, consultation following a diagnostic, and general consultation for systemic or preventive reason (Table S2).

Statistical analysis

Descriptive analysis of 4-HCS overall and subscale scores, consultation length, gender and occupation was performed. Normally distributed data were presented by mean and standard deviation (SD) and skewed distributions by the median and interquartile range (IQR). Binary and

categorical variables will be presented using counts and percentages.

Analyses were conducted for all participants, for medical students and for GP, HP and midwives combined (hereafter referred to medical practitioners) to account for the time limit on medical students' consultation.

We performed a linear regression analysis for continuous dependent variable with the 4-HCS score as dependent variable and length consultation and gender for explanatory variables.

Bellier et al. (36) identified gender as a confounder and we consider the reason for encounter as well, thus univariate analysis were subsequently adjusted on sex and reason for encounter.

We used multivariable fractional polynomial method for the regression model to investigate the relationship between the 4-HCS score and the consultation length. To compare the fits of these different models, we used an F-test, using the Anova function. The resulting curve was segmented into multiples segments using R software "segmented" package (version 1.4). Analyses were carried out on R software (version 4.0.5) and Stata (version 14.0).

Ethical considerations

After receiving appropriate written information, each participant signed an

individual consent form regarding research participation and image rights.

The study protocol was approved on October 21, 2020 by the CECIC Rhône-Alpes Auvergne, Clermont-Ferrand, France (IRB 5891). All data were de-identified before analysis.

3. Results

Sample Characteristics

Initially, 633 medical students recorded consultations with standardized patients, for 633 consultations, 10 GP registered 80 consultations, 26 HP registered 145 consultations and 12 midwives registered 58 consultations. We had to exclude 22 consultations (of 2 midwives) for no information on consultation length (Figure1).

The median consultation length, all groups combined, was 10.5 minutes [7.8, 15.8] and the median 4-HCS was 87.0 [74.0, 98.9]. Within subgroups, the 4-HCS score ranged from 84.0 [71.3, 95.0], for medical students, to 103.0 [91.5, 110.0] for midwives. The same distribution was found in all habits, except for Habit 2 (Table 1).

Primary outcome

For all participants and medical practitioners, the -2 power fractional polynomial fitted best, whereas the -1 power fitted best for medical students. Polynomial regression fitted statistically

better than linear regression ($p < 0.001$) for overall, medical students and medical practitioners (Figure 2).

Consultation length was also significantly associated ($p < 0.01$) with better 4-HCS in all analyses, and increase with the consultation length. Multivariate analysis had the similar results (Table 2).

The fractional polynomials for medical students and medical practitioners were segmented into two straight lines. For medical students and up to 6.85 minutes of consultation, the 4-HCS increase is 7.68 per minute, then 1.24 per minute. For medical practitioners, this increase is 3.15 per minutes up to 13.40 minutes of consultation, then 0.14 per minute (Figure 3).

Secondary outcomes

Women had significantly ($p=0.001$) better 4-HCS than men: 87 [76-98] versus 85 [67-97], without difference in consultation length (Table 4). The same results were found in all the 4-HCS dimensions, in overall and subgroups analyses (Table S1).

Surgeons and physicians had no difference on 4-HCS but consultations were longer for physicians ($p=0.002$) (Table 3).

ICPC-3, at the body system level, is defined by 19 different codes representing 1 to 24 consultations. The median consultations length ranged from 4.78 min for diagnoses and diseases of digestive system to 67.78 min [6.03] for general diagnoses and diseases. The 4-HCS score ranged from 67.00 for diagnoses and diseases of digestive system to 110.00 min for symptoms, complaints and abnormal findings of musculoskeletal system (Table 4).

ICPC-3, at the 3-category system level, is defined by 3 different codes representing 17 to 66 consultations. The median consultations length ranged from 26.13 [15.81] for consultation for systematic or preventive reason to 30.55 [26.14] for consultation following a diagnostic. The 4-HCS score ranged from 96.00 [18.00] for consultation for a symptom to 97.00 for the others (Table 5).

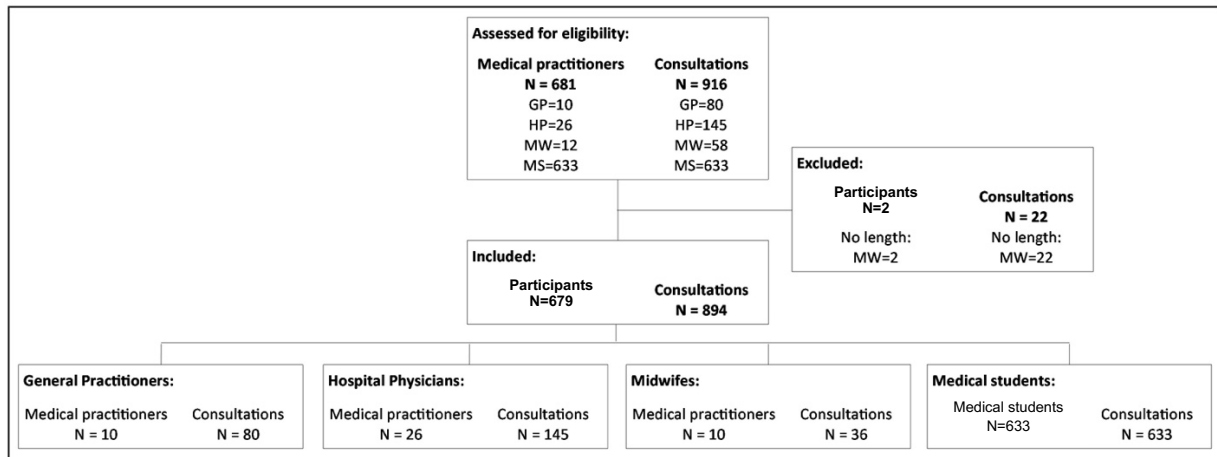


Figure 1. Flow chart of the study participants.

GP: General Practitioners, HP: Hospital Physicians, MW: Midwives, MS: Medical Students.

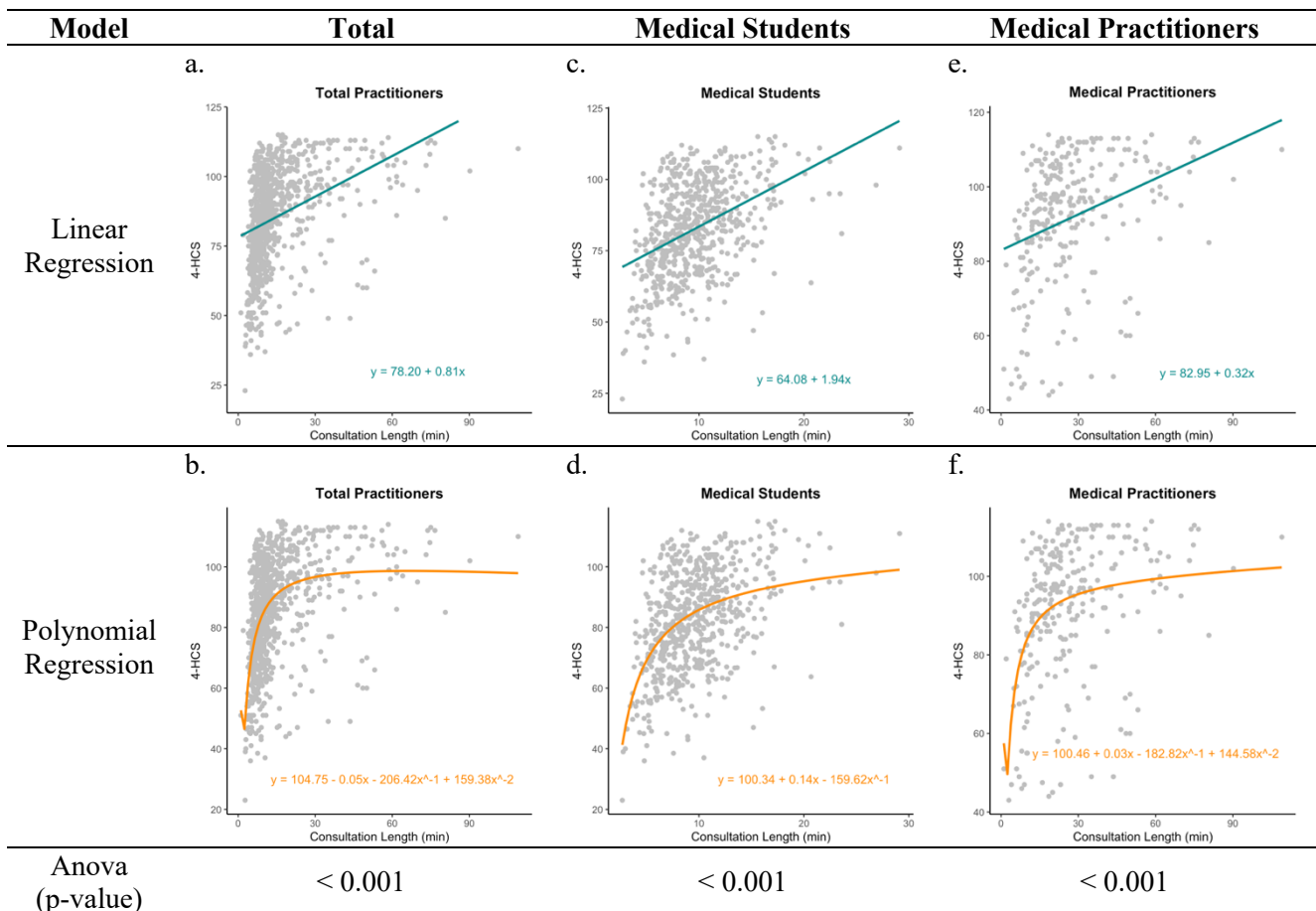


Figure 2. Graphs of the 4HCS score according to the consultation length:

- For the overall population using a linear regression model,
 - For the overall population according to a fractional polynomial model,
 - For medical students using a linear regression model,
 - For medical students using a fractional polynomial model,
 - For physicians and midwives using a linear regression model,
 - For physicians and midwives using a fractional polynomial model.
- Linear regression compared to polynomial regression with anova analysis.

Variable	Subgroup	
Consultation length in minutes (median [IQR])	General Practitioners	14.0 [10.0, 19.0]
	Hospital Physicians	29.8 [20.1, 43.6]
	Midwives	24.7 [20.2, 29.5]
	Medical Practitioners*	23.0 [14.8, 35.0]
	Medical Students	9.2 [7.2, 11.5]
	Total	10.5 [7.8, 15.8]
4-HCS total score (median [IQR])	General Practitioners	88.8 [76.8, 95.6]
	Hospital Physicians	97.0 [88.0, 106.0]
	Midwives	103.0 [91.5, 110.0]
	Medical Practitioners*	96.0 [85.0, 105.0]
	Medical Students	84.0 [71.3, 95.0]
	Total	87.0 [74.0, 98.9]
Habit1 score (median [IQR])	General Practitioners	22.3 [17.4, 25.1]
	Hospital Physicians	25.0 [22.0, 28.0]
	Midwives	28.0 [25.8, 30.0]
	Medical Practitioners*	25.0 [21.0, 28.0]
	Medical Students	20.5 [17.5, 24.0]
	Total	21.6 [18.0, 25.0]
Habit2 score (median [IQR])	General Practitioners	10.5 [8.5, 12.1]
	Hospital Physicians	13.0 [11.0, 14.0]
	Midwives	13.0 [10.0, 14.0]
	Medical Practitioners*	12.0 [10.0, 14.0]
	Medical Students	11.8 [9.7, 13.0]
	Total	12.0 [9.8, 14.0]
Habit3 score (median [IQR])	General Practitioners	14.5 [12.5, 16.5]
	Hospital Physicians	15.0 [14.0, 17.0]
	Midwives	17.0 [14.0, 20.0]
	Medical Practitioners*	15.0 [13.0, 17.0]
	Medical Students	14.0 [12.0, 17.0]
	Total	14.5 [12.0, 17.0]
Habit4 score (median [IQR])	General Practitioners	41.0 [34.9, 43.5]
	Hospital Physicians	45.0 [40.0, 49.0]
	Midwives	44.5 [40.0, 47.3]
	Medical Practitioners*	43.5 [39.0, 47.0]
	Medical Students	37.0 [31.0, 43.0]
	Total	39.0 [32.3, 45.0]
Gender (% of women)	General Practitioners	39 (48.8%)
	Hospital Physicians	82 (56.6%)
	Midwives	36 (100%)
	Medical Practitioners*	157 (60.2%)
	Medical Students	398 (62.9%)
	Total	555 (62.1%)
Specialty for Hospital Physicians (%)	Surgeon	19 (13.1%)
	Physician	126 (86.9%)

Table 1. Characteristics of the participants according to subgroup

*Medical practitioners correspond to General Practitioners, Hospital Physicians and Midwives

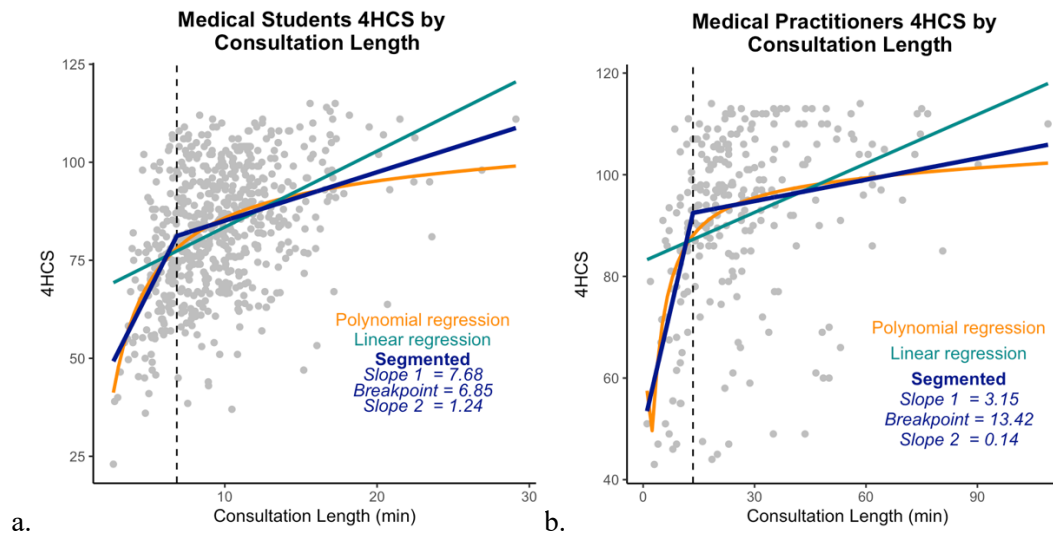


Figure 3. 4-HCS by consultation length, polynomial regression, linear regression and segmented lines for: a. medical students, b. medical practitioners.

Groups	Consultation length (min) Median [IQR]	Total Score Median [IQR]	Univariate β coeff (IC95%)	p-value	Multivariate β coeff (IC95%)	p-value				
Total Practitioners	1st quartile [1.05,7.77]	74.25 [63.88, 87.00]	0.49 (0.40, 0.57)	<0.00 1	0.48 (0.40, 0.56)	<0.001				
	2 nd quartile (7.77,10.5]	83.50 [73.81, 94.00]								
	3rd quartile (10.5,15.8]	90.00 [79.83, 99.00]								
	4th quartile. (15.8,109]	98.00 [90.00, 106.12]								
	1st quartile [1.05,14.8]	85.25 [69.62, 93.38]								
Medical Practitioners	2 nd quartile (14.8,23]	97.00 [89.12, 103.25]	0.32 (0.21, 0.43)	<0.00 1	0.31 (0.20, 0.42)	<0.001				
	3rd quartile. (23,35]	97.00 [89.00, 105.50]								
	4th quartile (35,109]	102.00 [92.00, 110.00]								
	1st quartile [2.68,7.17]	72.00 [62.00, 83.38]								
	2 nd quartile (7.17,9.2]	83.00 [73.44, 94.00]								
Medical Students	3rd quartile (9.2,11.5]	86.25 [75.06, 96.00]	1.93 (1.62, 2.26)	<0.00 1	1.94 (1.62, 2.26)	<0.001				
	4th quartile (11.5,29.1]	93.00 [83.12, 102.00]								
	1st quartile [2.0,20.0]	93.00 [81.00, 104.50]								
	2 nd quartile (20.3,28.6]	94.00 [84.00, 103.25]					0.22 (0.07, 0.37)	0.005	0.30 (0.12, 0.48)*	0.001*
	3rd quartile (28.6,43.5]	97.00 [90.00, 105.50]								
4th quartile (43.5,109]	106.00 [96.75, 110.50]									
3rd quartile (28.6,43.5]	97.00 [90.00, 105.50]	0.22 (0.07, 0.37)**	0.005**							
4th quartile (43.5,109]	106.00 [96.75, 110.50]									

Table 2. Univariate 4-HCS analysis and multivariate analysis on gender and consultation length.

*Multivariate analysis on gender, consultation length and ICPC-3 at the body system level.

** Multivariate analysis on gender, consultation length and ICPC-3 at the 3-category system level.

Variable	Subgroup	Total Score Median [IQR]	p-value	Consultation Length (min) Median [IQR]	p-value	
Gender	General Practitioners	Women (n=39)	93.0 [86.0, 97.0]	<0.001	17.0 [12.5, 22.5]	<0.001
		Men (n=41)	78.0 [61.5, 92.5]		11.0 [8.0, 16.0]	
	Hospital Practitioners	Women (n=82)	99.0 [89.3, 107.8]	0.125	30.9 [19.6, 45.1]	0.803
		Men(n=63)	97.0 [85.5, 103.0]		29.6 [22.8, 43.1]	
	Medical Students	Women (n=398)	85.0 [74.0, 96.0]	0.004	9.2 [7.2, 11.3]	0.893
Men (n=235)		82.0 [66.0, 94.0]	9.4 [7.0, 11.8]			
Total	Women (n=555)	87.0 [76.0, 98.0]	0.001	10.1 [7.7, 14.7]	0.567	
	Men (n=339)	85.0 [67.0, 97.0]		10.7 [7.6, 15.2]		
Specialty	Hospital Practitioners	Surgeon (n=19)	101.0 [86.5, 105.0]	0.870	19.7 [15.2, 27.1]	0.002
		Physician (n=126)	97.0 [88.3, 107.0]		31.9 [21.5, 46.2]	
Year	Medical student	2017-18 (n=197)	75.7 [67.0, 84.0]	<0.001	9.2 [7.4, 11.3]	0.072
		2018-19 (n=180)	88.0 [76.0, 99.0]		8.7 [6.9, 11.2]	
		2019-20 (n=161)	87.0 [74.0, 99.0]		9.6 [7.0, 11.6]	
		2020-21 (n=95)	93.0 [77.5, 101.5]		10.1 [7.4, 11.8]	

Table 3. Characteristics of the participants according to total score and consultation length

4. Discussion & Conclusion

Results interpretation

The median consultation time for medical practitioners was 23.0 minutes and the 4-HCS was 96.0. For medical students, the median consultation time was 9.2 minutes and the 4-HCS was 84.0. The time limit for medical students decreases the consultation length and may explain the decrease in 4-HCS.

For medical practitioners, consultation length was heterogeneous, ranging from 14.0 for GP to 29.8 minutes for HP. These differences could be explained by the diversity of reasons for encounter: consulting for the first time or for follow-up

does not require the same consultation length, as does a consultation to announce a cancer diagnosis. As the complexity of the consultation increases, so does the time required. For example, it has been shown that a consultation for a second opinion takes longer than the first (37).

It is also possible that the different remuneration model for these professionals may influence the length of the consultation. Since general practitioners are paid on a fee-for-service basis, they have an incentive to increase the number of consultations. This is consistent with our results; GP had the shortest consultation.

Group by body system and symptom or diagnoses	Code	Count (n=111)	Consultation Length (min) Median [IQR]	Total Score Median [IQR]
Family planning	AF	3	32.6 [5.9]	96.0 [21.0]
General and routine examination	AG	3	25.2 [11.3]	108.0 [4.0]
Visit for other reasons	AR	22	25.0 [16.2]	97.0 [13.8]
General symptoms, complaints and abnormal findings	AS	6	51.7 [51.6]	94.5 [20.0]
General diagnoses and diseases	AD	2	67.8 [6.0]	105.0 [7.0]
Symptoms, complaints and abnormal findings of blood, blood-forming organs and immune system	BS	2	17.9 [12.6]	101.5 [10.5]
Diagnoses and diseases of blood, blood-forming organs and immune system	BD	24	32.3 [24.8]	96.00 [16.8]
Diagnoses and diseases of digestive system	DD	1	4.8	67.0
Symptoms, complaints and abnormal findings of eye	FS	1	35.3	90.0
Symptoms, complaints and abnormal findings of genital system	GS	5	28.2 [13.5]	96.0 [13.0]
Symptoms, complaints and abnormal findings of genital system	GD	5	20.1 [14.1]	109.0 [6.0]
Diagnoses and diseases of circulatory system	KD	14	36.2 [22.8]	94.0 [36.0]
Symptoms, complaints and abnormal findings of musculoskeletal system	LS	1	43.4	110.0
Diagnoses and diseases of musculoskeletal system	LD	8	17.4 [15.7]	104.0 [13.8]
Diagnoses and diseases of neurological system	ND	3	15.1 [7.3]	93.0 [12.5]
Symptoms, complaints and abnormal findings of respiratory system	RS	2	23.8 [3.9]	98.0 [6.0]
Diagnoses and diseases of skin	SD	2	11.5 [3.9]	83.5 [2.5]
Diagnoses and diseases of endocrine, metabolic and nutritional system	TD	3	31.3 [11.5]	88.00 [25.0]
Symptoms, complaints and abnormal findings of urinary system	UD	4	22.9 [10.9]	105.0 [11.0]

Table 4. ICPC-3 characteristic at the body system level

Group by	Code	Count (n=11)	Consultation Length Median [IQR]	Total Score Median [IQR]
consultation following a diagnostic	D	66	30.6 [26.1]	97.0 [25.0]
consultation for a symptom	S	17	28.8 [11.2]	96.0 [18.0]
consultation for systemic or preventive reason	G	28	26.1 [15.8]	97.0 [15.3]

Table 5. ICPC-3 characteristic at the 3-category system level

The duration of the consultation was significantly associated with the interpersonal skills score since the score increased by 0.49 points after each additional minute of consultation ($\beta=0.49$, $p<0.001$) and by 0.32 points among medical practitioners ($\beta=0.32$, $p<0.001$). This is how we can explain the fact that hospital

physicians have a higher score since they spent more time in consultation compared to general practitioners and students. Nevertheless, midwives spent less time in consultation than hospital physicians but seems to have better 4-HCS, which could be explain by a different kind of formation and different reasons for encounter.

Furthermore, we can observe a threshold effect at which the level of interpersonal skills becomes satisfactory and increases less rapidly with the increase in the duration of the consultation. Thus, for medical practitioners, spending more than 13.42 minutes in consultation gives a median score of 97.2 [89.9, 106.0] out of 115. For students, after 6.85 minutes of consultation, 4-HCS the median was of 86.9 [76.0, 98.0] out of 115. Indeed, from the fractional polynomial regression model which was a more relevant model to understand and describe the evolution of relational skills as a function of consultation duration, we can adjust our statement. For students, the gain in score was greater: the score increased by 7.68 points up to 6.85 minutes and then only by 1.24 points. For medical practitioners, the score increased by 3.15 points after each additional minute of consultation until 13.42 minutes, and then the score gain was only 0.14 points. Thus, it is the first few minutes that appear to be most important in increasing the level of interpersonal skills. Finally, good interpersonal skills and average consultation time are not mutually exclusive. Moreover, like Licqurisk et al. have shown (38), the use of tools to facilitate patient-physician communication makes it possible not to increase consultation time and may improve symptom detection or reduce overall pain symptoms.

However, physician-patient communication must also adapt to new challenges, including the emergence of video-consultation in connection with the COVID-19 crisis. Hammersley et al. (39) have shown the impact on the reduction of the consultation time and quality of consultations compared to a face-to-face consultation.

Our results shown that women had better 4-HCS ($\beta=-5.34$, $p<0.001$), these results are in agreement with the literature which shows that women have better empathy towards their patients than men, as well as better non-verbal communication (36). Our findings reinforce this idea since in dimensions 2 and 3, which are respectively "Obtaining the patient's point of view" and "Showing empathy", women obtained better results.

Strength of the study

Our cohort exceeds 500 data, which is rare in the literature. Video combined with audio is also an advantage over audio alone since it collects more information during the interview, particularly concerning non-verbal communication (32). We recruited medical practitioners of different ages, in several medical and surgical specialties, with several statuses ranging from students to university professors, and in urban, semi-rural or rural medicine. This provided a

large population study to represent medicine practice.

Study limitations and bias

In the students' exercise, instruction was given not to exceed ten minutes of simulated consultation. However, the time was not strictly controlled, so the students could exceed this limit. Moreover, the consultation was limited to the questioning, without any physical examination of the patient, thus mechanically shortened the consultation time. In addition, the 4-HCS does not take into account limit time and reason for encounter on interpersonal skills. However, a comparison of student consultations with limited time and physician consultations did not show a significant difference, supporting the idea that this exercise enables to assess students' interpersonal skills.

In addition, in the data collection of general practitioners, the majority of doctors were involved in the training of young doctors (university lecturers, internship supervisors, young doctors leaving the internship) because of the difficulties of recruitment outside this network. We can assume that these doctors were already aware of the quality of communication within their practice. On the other hand, the physicians were all volunteers, and therefore probably more aware of the problem of relationships.

Technically, the study had its limitations because the consultation did not always take place completely in the office in the field of vision of the camera, as the reception and the end of the consultation could take place on the doorstep or in the waiting room.

Another interesting point is that the general practitioners were not limited to receiving only primary care patients, so their patients were more heterogeneous than the patients of the students and hospital doctors, who were all primary care patients.

Conclusion

We have shown that the length of the consultation has a strong influence on the level of interpersonal skills of the physician. Indeed, among medical students and medical practitioners, the duration of the consultation significantly increased the score of the 4-HCS, especially in the first minutes. Each additional minute of consultation increased the score by 7.68 points up to 6.85 minutes of consultation for students and 3.15 points up to 13.42 minutes of consultation for physicians, after which the gain in score became small. Thus, a relationally effective medical consultation need not exceed the average consultation time. Furthermore, women scored higher than men, and did better on the dimensions of empathy and obtaining the patient's perspective. It therefore seems necessary to develop interpersonal skills in initial and

continuing training, particularly for men.
The 4-HCS could be a training aid in this
context.

References

1. Ong LML, de Haes JCJM, Hoos AM, Lammes FB. Doctor-patient communication: A review of the literature. *Soc Sci Med.* 1 avr 1995;40(7):903-18.
2. What is the most effective way to maintain weight loss in adults? | The BMJ [Internet]. [cité 7 déc 2021]. Disponible sur: <https://www.bmj.com/content/343/bmj.d8042>
3. Boissy A, Windover AK, Bokar D, Karafa M, Neuendorf K, Frankel RM, et al. Communication Skills Training for Physicians Improves Patient Satisfaction. *J Gen Intern Med.* juill 2016;31(7):755-61.
4. Haskard Zolnieriek KB, DiMatteo MR. Physician Communication and Patient Adherence to Treatment: A Meta-Analysis. *Med Care.* août 2009;47(8):826-34.
5. Kelley JM, Kraft-Todd G, Schapira L, Kossowsky J, Riess H. The Influence of the Patient-Clinician Relationship on Healthcare Outcomes: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *PLOS ONE.* 9 avr 2014;9(4):e94207.
6. Mercer SW, Fitzpatrick B, Guthrie B, Fenwick E, Grieve E, Lawson K, et al. The CARE Plus study - a whole-system intervention to improve quality of life of primary care patients with multimorbidity in areas of high socioeconomic deprivation: exploratory cluster randomised controlled trial and cost-utility analysis. *BMC Med.* 22 juin 2016;14(1):88.
7. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician-patient communication. The relationship with malpractice claims among primary care physicians and surgeons. *JAMA.* 19 févr 1997;277(7):553-9.
8. Osborn R, Moulds D, Schneider EC, Doty MM, Squires D, Sarnak DO. Primary Care Physicians In Ten Countries Report Challenges Caring For Patients With Complex Health Needs. *Health Aff Proj Hope.* déc 2015;34(12):2104-12.
9. Wilson A, McDonald P, Hayes L, Cooney J. Longer booking intervals in general practice: effects on doctors' stress and arousal. *Br J Gen Pract J R Coll Gen Pract.* mai 1991;41(346):184-7.
10. Wang Q, Adhikari SP, Wu Y, Sunil TS, Mao Y, Ye R, et al. Consultation length, process quality and diagnosis quality of primary care in rural China: A cross-sectional standardized patient study. *Patient Educ Couns.* 8 août 2021;S0738-3991(21)00545-0.

11. Mercer SW, Fitzpatrick B, Gourlay G, Vojt G, McConnachie A, Watt GC. More time for complex consultations in a high-deprivation practice is associated with increased patient enablement. *Br J Gen Pract.* 1 déc 2007;57(545):960-6.
12. Wilson A, Childs S. The relationship between consultation length, process and outcomes in general practice: a systematic review. *Br J Gen Pract J R Coll Gen Pract.* déc 2002;52(485):1012-20.
13. Rodriguez KL, Bayliss NK, Alexander SC, Jeffreys AS, Olsen MK, Pollak KI, et al. Effect of patient and patient-oncologist relationship characteristics on communication about health-related quality of life. *Psychooncology.* sept 2011;20(9):935-42.
14. Valverde Bolívar FJ, Pedregal González M, Moreno Martos H, Cózar García I, Torío Durántez J. Communication with patients and the duration of family medicine consultations. *Aten Primaria.* déc 2018;50(10):621-8.
15. Lemon TI, Smith RH. Consultation Content not Consultation Length Improves Patient Satisfaction. *J Fam Med Prim Care.* 2014;3(4):333-9.
16. Elmore N, Burt J, Abel G, Maratos FA, Montague J, Campbell J, et al. Investigating the relationship between consultation length and patient experience: a cross-sectional study in primary care. *Br J Gen Pract J R Coll Gen Pract.* déc 2016;66(653):e896-903.
17. Effects of interventions aimed at changing the length of primary care physicians' consultation - Wilson, AD - 2006 | Cochrane Library [Internet]. [cité 7 déc 2021]. Disponible sur: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD003540.pub2/full>
18. Irving G, Neves AL, Dambha-Miller H, Oishi A, Tagashira H, Verho A, et al. International variations in primary care physician consultation time: a systematic review of 67 countries. *BMJ Open.* 8 nov 2017;7(10):e017902.
19. Jin G, Zhao Y, Chen C, Wang W, Du J, Lu X. The length and content of general practice consultation in two urban districts of Beijing: a preliminary observation study. *PloS One.* 1 janv 2015;10(8):e0135121.
20. Orton PK, Pereira Gray D. Factors influencing consultation length in general/family practice. *Fam Pract.* oct 2016;33(5):529-34.
21. Roter D, Lipkin M, Korsgaard A. Sex differences in patients' and physicians' communication during primary care medical visits. *Med Care.* nov 1991;29(11):1083-93.
22. Westcott R. The length of consultations in general practice. *J R Coll Gen Pract.* sept 1977;27(182):552-5.
23. Wiggers JH, Sanson-Fisher R. Duration of general practice consultations: association with patient occupational and educational status. *Soc Sci Med* 1982. avr 1997;44(7):925-34.

24. Blankfield RP, Goodwin M, Jaén CR, Stange KC. Addressing the unique challenges of inner-city practice: a direct observation study of inner-city, rural, and suburban family practices. *J Urban Health Bull N Y Acad Med.* juin 2002;79(2):173-85.
25. Jefferson L, Bloor K, Birks Y, Hewitt C, Bland M. Effect of physicians' gender on communication and consultation length: a systematic review and meta-analysis. *J Health Serv Res Policy.* oct 2013;18(4):242-8.
26. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: a meta-analytic review. *JAMA.* 14 août 2002;288(6):756-64.
27. Nicolai J, Demmel R. The impact of gender stereotypes on the evaluation of general practitioners' communication skills: an experimental study using transcripts of physician-patient encounters. *Patient Educ Couns.* déc 2007;69(1-3):200-5.
28. Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J. Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills. *PLOS ONE.* 16 avr 2020;15(4):e0230672.
29. Blanc G, Bellier A, Di Marco L. Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en maïeutique. *Sages-Femmes.* 1 juill 2020;19(4):50-4.
30. Bellier A, Labarère J, Putkaradze Z, Cavalie G, Carras S, Pelen F, et al. Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): protocol for a randomised controlled trial. *BMJ Open.* 15 févr 2022;12(2):e051600.
31. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet Lond Engl.* 20 oct 2007;370(9596):1453-7.
32. Asan O, Montague E. Using video-based observation research methods in primary care health encounters to evaluate complex interactions. *Inform Prim Care.* 2014;21(4):161-70.
33. ICPC-3 [Internet]. [cité 21 févr 2022]. Disponible sur: <https://www.icpc-3.info/>
34. Krupat E, Frankel R, Stein T, Irish J. The Four Habits Coding Scheme: Validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns.* 1 juill 2006;62(1):38-45.
35. Enhancing clinician communication skills in a large healthcare organization: A longitudinal case study - ScienceDirect [Internet]. [cité 7 déc 2021]. Disponible sur:

<https://www.sciencedirect.com/science/article/abs/pii/S0738399105000133?via%3Dihub>

36. Bellier A. Factors associated with interpersonal skills in medical students training for standardized medical consultations. :57.
37. Lehmann V, Smets EMA, de Jong M, de Vos FYF, Stouthard JM, Hillen MA. Patient-provider communication during second opinion consultations in oncology. *Patient Educ Couns.* oct 2021;104(10):2490-7.
38. Licqurish SM, Cook OY, Pattuwage LP, Saunders C, Jefford M, Koczwara B, et al. Tools to facilitate communication during physician-patient consultations in cancer care: An overview of systematic reviews. *CA Cancer J Clin.* nov 2019;69(6):497-520.
39. Hammersley V, Donaghy E, Parker R, McNeilly H, Atherton H, Bikker A, et al. Comparing the content and quality of video, telephone, and face-to-face consultations: a non-randomised, quasi-experimental, exploratory study in UK primary care. *Br J Gen Pract J R Coll Gen Pract.* sept 2019;69(686):e595-604.

CHAPITRE 3

Développement des compétences relationnelles par des programmes pédagogiques dédiés

4. CHAPITRE 3 : Développement des compétences relationnelles par des formations dédiées en éducation médicale

Après avoir étudié les déterminants des compétences relationnelles, grâce à l'échelle 4-HCS, nous avons cherché à déterminer par quels moyens nous pouvions agir sur ces compétences. Nous avons donc débuté par l'évaluation de programmes pédagogiques existants en formation initiale dans un premier temps, puis par la mise en place d'un programme pédagogique multifacette dédié en formation continue.

4.1. Partie 7) Évaluation des dispositifs d'amélioration des compétences relationnelles lors de la formation initiale

Article publié — Impact of physician–patient relationship training on medical students' interpersonal skills during simulated medical consultations. A cross-sectional study

Dans cet article publié le 21 février 2022 dans la revue BMC Medical Education, nous avons cherché à évaluer l'impact de plusieurs formations dispensées au cours du premier cycle des études de médecine sur les compétences relationnelles, évaluées lors de simulations avec patients standardisés en quatrième année de médecine.

Contexte

En médecine, l'approche centrée sur le patient repose sur des compétences relationnelles, notamment la communication, la structuration de l'entretien médical et l'empathie, qui ont un impact sur les relations relationnelles des professionnels de la santé et la qualité des soins. Des formations sur ce thème sont donc développées dans les universités. Nous avons émis l'hypothèse que des formations spécifiques à la relation médecin-patient pourraient améliorer les compétences relationnelles des étudiants en médecine lors de consultations simulées et la satisfaction immédiate des patients standardisés.

Méthodes

Cette étude transversale a recruté des étudiants en médecine de quatrième année qui ont participé à une session de consultation médicale simulée avec des patients standardisés. L'évaluation des compétences relationnelles a été réalisée à l'aide de l'échelle 4-HCS, produisant un score synthétique sur 115 points utilisé comme critère d'évaluation principal. Certains étudiants ont bénéficié des formations proposées par l'université ou par d'autres organisations, principalement basées sur la communication, l'écoute active ou l'approche centrée sur le patient. Une comparaison a été faite avec les étudiants de la même promotion qui n'avaient reçu aucune formation spécifique aux compétences relationnelles.

Résultats

L'analyse du critère principal a montré une différence de 5 points sur le score 4-HCS entre le groupe d'étudiants ayant suivi au moins une formation et les autres ($p = 0,001$). Cette différence était encore plus marquée lorsque les étudiants avaient suivi plusieurs formations, jusqu'à 14 points de plus au score 4-HCS avec trois formations ($p = 0,001$), chacune ayant des résultats positifs dans différents domaines de la relation de soins.

Conclusions

La formation à la relation médecin-patient actuellement dispensée en formation initiale semble efficace pour améliorer les compétences relationnelles. Une répétition de cette formation est nécessaire pour augmenter son impact.

RESEARCH

Open Access



Impact of physician–patient relationship training on medical students' interpersonal skills during simulated medical consultations: a cross-sectional study

Lucie Bosméan¹, Philippe Chaffanjon² and Alexandre Bellier^{2,3,4*} 

Abstract

Background: In medicine, the patient-centered approach is based on interpersonal skills, including communication, structuring the medical interview, and empathy, which have an impact on health professionals' interpersonal relationships and the quality of care. Training courses on this issue are therefore being developed in universities. We hypothesized that specific training courses in the physician–patient relationship could improve interpersonal skills among medical students during simulated consultations and the immediate satisfaction of standardized patients.

Methods: This cross-sectional study enrolled fourth-year medical students who participated in a simulated medical consultation session with standardized patients. The evaluation of interpersonal skills was carried out using the Four Habits Coding Scheme, producing a synthetic score out of 115 points used as the primary endpoint. Some students benefited from the training courses offered by the university or by other organizations, mainly based on communication, active listening, or patient-centered approach. A comparison was made with students from the same graduating class who had not received any training.

Results: The analysis of the primary endpoint showed a difference of 5 points between the group of students who had attended at least one training course and those who did not ($p = 0.001$). This difference was even more marked when the students had completed several training courses, up to 14 points higher with three training courses ($p = 0.001$), each with positive results in different areas of the care relationship.

Conclusions: Physician–patient relationship training currently provided in initial education appears to be effective in improving interpersonal skills. A repetition of this training is necessary to increase its impact.

Keywords: Clinical competence, Health education, Interpersonal skills, Medical students, Simulation

Background

In medicine, the patient-centered approach is based on interpersonal skills, including communication, structuring the medical interview, and empathy, which have an

impact on health professionals' interpersonal relationships and the quality of care [1, 2]. Interpersonal skills are defined as the presence of verbal and nonverbal behaviors in the context of personal interactions with the patient or the patient's family [3]. French universities are starting to offer training courses on this specific issue. Indeed, interpersonal skills are one of the essential skills to be taught in medical curricula and are among the most appreciated by patients [4–6]. They increase the quality of care and

*Correspondence: abellier@chu-grenoble.fr

⁴ Computational and Mathematical Biology Team, TIMC UMR 5525, CNRS, University of Grenoble Alpes, Grenoble, France

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

help decrease human or economic costs due to adverse events [7]. Finally, it has been shown that communication skills tend to decline over time unless they are regularly recalled and practiced [8]; the same applies to empathy, which also decreases over time – a finding that has been validated since 2010 [4].

In the interest of the patient and the quality of care, it is therefore essential to consider whether training can help halt this decline, and to reflect on the personal benefit to the practitioner. Several studies have indicated that communication skills can be taught and learned in both simulated and real clinical settings [9–11]. It was shown that short-term training focusing on interpersonal skills can lead to a significant change in behavior ($p=0.010$) and long-term self-efficacy ($p=0.042$) for senior physicians [12]. In a review of the literature, it was reported that educational interventions can also be effective in maintaining and reinforcing empathy among medical students from the beginning of training, with a mean effect of 0.23 on the Medical Education Research Study Quality Instrument score [13].

It is sometimes difficult to demonstrate the effectiveness of these programs on physician competencies, and generally only one component of the physician–patient relationship is assessed, such as empathy [14] or communication [9]. Another gap highlighted by a recent study is the difficulty in demonstrating a transfer of skills from physician to patient [15]. Few studies have been able to show an impact on the patient [16, 17], and only one randomized trial [17] focused both on the impact of training on interpersonal skills for the hospital as a whole and on the impact for the patient in terms of immediate satisfaction. Furthermore, the population studied was largely physicians, more often at the hospital, and less often medical students. Finally, the retention over time of the skills acquired is a crucial point. It seems that the effect of training, in this case a multifaceted program inspired by the Kaiser Permanente [12], on interpersonal skills may persist in the long term. However, few studies have measured the medium- or long-term retention of skills acquired through training, and to our knowledge, no studies have examined the long-term retention of interpersonal skills of medical students in training targeted at interpersonal skills.

In order to carry out such evaluations in medical consultation, it is necessary to have standardized, valid, and reliable instruments [5]. On the basis of its satisfactory psychometric properties, several entities in numerous countries use the Four Habits Coding Scheme (4-HCS) [1] to assess interpersonal skills (Empathy, Interview Structuring, and Communication) [16, 17]. This scale has four sub-sections: Involvement from the beginning (Habit 1), Getting the patient's point of view (Habit 2),

Demonstrating empathy (Habit 3), and Involvement until the end (Habit 4). For each of the 23 items, the rating uses a 5-point Likert scale. By adding up the points, the scale produces a composite score from 23 to 115 points, as well as specific scores for each dimension. The 4-HCS has been used with thousands of physicians worldwide [12, 17–20] and a French translation has been validated [5].

The main objective of our study was therefore to measure the impact that training in physician–patient relationships can have on the interpersonal skills of medical students during simulated consultations. The secondary objectives were to compare the effectiveness of the different training courses on improving the students' interpersonal skills and to evaluate the immediate satisfaction of standardized patients at the end of the consultations.

Methods

Study design

We conducted a cross-sectional study in the field of medical education, with a prospective recruitment design. The study took place from October 2019 to January 2020 at Grenoble Alpes University Hospital. We followed the international STROBE guidelines for this study [21].

Participants

Fourth-year medical students enrolled in the Faculty of Medicine of the University of Grenoble Alpes during the 2019–2020 academic year were eligible to participate in the study. Students were required to participate in the practice as part of their curriculum. Each medical student participated in a single simulated consultation. At the time of the evaluation, each student had the same clinical background. We excluded students who were not available at the time of their convocation.

Data collection

The requested exercise consisted in carrying out a consultation limited to questioning with the participation of actors from the university's Department of Performing Arts, as standardized patients [2]. The objective was to place the medical student in a realistic clinical situation (consultation for common pathologies, without emergency criteria and accessible to outpatients, as in a general medicine consultation) with typical stereotyped patients. This exercise was the first simulated consultation for these students. All simulated consultations were video-recorded, and the videos were stored in a database on a secure cloud for later remote access [2]. The video made it possible to perform one or more evaluations after the consultation, without the need for the physical presence of an evaluator during the consultation, which could affect the medical student's behavior [22]. Each

consultation was evaluated by one of three physicians with experience in interpersonal skills assessment. These physicians evaluated the videos independently of each other during the same period. Evaluators completed an online form in which all items on the scale were mandatory, thus there could be no missing data.

The evaluation of interpersonal skills was made using the 4-HCS [2] scale, which resulted in a total score and four sub-scores corresponding to each dimension (Habit 1, Habit 2, Habit 3, and Habit 4). The 4-HCS score was the primary endpoint of the study. The evaluation of the videos and the calculation of the 4-HCS score were performed blind to the type of training undertaken by the student. Performing Arts students were asked to complete an evaluation of service consumers and health systems using a visual analog scale (VAS) rated from 0 to 10: "Using this visual analog scale, where this end represents the worst possible doctor and the other end represents the best possible doctor, please rate this doctor by placing a cross corresponding to how you feel..." This scale was used to measure overall patient satisfaction [12], which was used as a secondary endpoint for evaluating interpersonal skills.

Training courses

Our students were able to follow various short-term training courses on a voluntary basis to improve their interpersonal skills. Some of the courses were academic, others not, but we chose to evaluate all of them. Each course could be followed only once by a student. The first course, an academic training course in communication, was given by the Faculty of Medicine, in small groups (15 students) in the form of a 3-h interactive course: exchanges, role plays, and viewing of consultation examples. It dealt with ethics, the legislative framework of medical practice, and medical communication according to the Four Habits Model developed by the Kaiser Permanente Institute [16]. The course was given on an experimental basis 1 year before the simulated consultation exercise to approximately one third of the students in the class. The second course, peer training, was given by peers trained in active listening and medical communication, and took place in small groups (15 students) in the form of a 1-day interactive course that involved role-playing. It was based on Rogers's active listening, emotional intelligence, and different communication profiles/channels. It was given 1 year before the simulated consultation exercise. The third training course, given by an association, was carried out by a psychologist in small groups (15 students) in the form of a 4-day interactive course: exchanges, work on photography, and role-playing. It was based on active listening according to Rogers and was given 2 years before the simulated consultation

exercise. Finally, the fourth training set included training not mentioned above and was carried out by the students in a framework other than that of the university, having a link with medical communication. This included training in transactional analysis, bereavement support, medical communication courses in midwifery school, non-compulsory university teaching on speaking and listening in care, training in difficult announcements, training in team management, and training in support for children with disabilities.

Before the simulated consultations, the students were asked about their participation in one of these training courses conducted on a voluntary basis. All the training sessions were held more than 1 year before the simulated consultations and the evaluation focused on the long-term retention of the skills acquired. After giving information about the study and obtaining consent, the student volunteers were recruited and divided into two groups: a training group and a control group. In the training group, the students had partaken in at least one of the aforementioned training courses, while the students in the control group had not received any specific training in caregiver relationship skills. Thanks to the participation of all students, no data were missing.

Statistical methods

Calculating the number of participants required was based on the study by Gulbrandsen et al. [12], which described an improvement in the 4-HCS score from 58.8 to 62.9 points after 20 h of specific training (standard deviation=10). With an alpha risk of 5% and a power of 80%, the minimum number of participants was 95. Inclusion characteristics were described for the entire sample and for each training received. Univariate analyses were performed to compare student characteristics with the 4-HCS score using Student's *t* test or a Mann-Whitney test depending on the distribution of the quantitative variable, as well as an ANOVA or Kruskal-Wallis test and a Pearson or Spearman correlation test depending on the validity conditions. The primary endpoint was analyzed using a multivariate analysis with linear regression model. In the case of an association of less than 0.20 between a measured characteristic and the 4-HCS score, the analysis was performed by introducing this variable as a fixed-effect covariate into the model. This adjustment was carried out on the standardized patient criteria ($p=0.031$), on the day of the simulated consultation ($p=0.185$), and on the number of training sessions ($p=0.192$). The secondary endpoint was analyzed using linear regression models, following a similar strategy. The statistical significance level was set at 0.05 in the two-tailed situation for all comparison tests. No *p*-value adjustments were expected. Statistical analyses

were performed using the software RStudio (Version 1.0.143©).

Ethics

Students were given information about the research in person and were included only after signing an individual consent form and image rights. Ethics committee approval was granted on 22th February 2021, by the regional ethics committee: *Comité d’Ethique du Centre d’Investigation Clinique de Clermont-Ferrand* (IRB 00005891).

Results

Our sample consisted of 163 students who were invited to participate in the simulation of medical consultations. Two students were absent and no student refused to participate. Thus, we recorded and analyzed 161 videos of simulated consultations (Fig. 1). The average age of the students was 21 years and 69% were female (N=112). A total of 72 students (44.7%) participated in interpersonal skills trainings: 48 students (29.8%) participated in the academic training, 14 (8.6%) in the training by an association, 15 (9.3%) in the peer training, and eight (4.9%) in other training as previously defined. Overall, 61 students (37.8%) participated in only one type of training.

Analysis of the main outcome (i.e., 4-HCS synthetic score) showed a difference of 6.7 points between the group of students with at least one training course and those without, with an average of 85 points (SD=17.2) for all the students (Table 1). In multivariate analysis, the

difference between the two groups of students was statistically significant (p=0.001). It should be noted that this difference was even more marked when the students had completed several courses. Without training, students had a mean score of 82 points/115 (SD=18.7), whereas after three different training courses the mean score was 109 points/115 (SD=6.36). The difference in the 4-HCS score depending on the number of training courses taken was statistically significant (p<0.001).

When considering each training independently, there was a 5.1-point increase in the 4-HCS score for academic training (p=0.002), a 5.9-point increase for peer training (p=0.031), and an 8.9-point increase for other training (p=0.007) (Table 1).

Analyzing the 4-HCS score according to the 4 Habits, several training courses were found to have an impact on Habit 1 with a 0.7-point increase in the score on this dimension (p=0.047) for academic training and a 2.1-point increase (p=0.025) out of 30 points for peer training. Academic training increased the score for Habit 2 by 1.1 points (p=0.002) out of 15 points. For Habit 3, academic training increased by 1.2 points (p=0.004) and other training by 2.3 points out of 20 points (p=0.007). Finally, for Habit 4, the increase in score was 2.2 points (p=0.005) for academic training and 2.7 points (p=0.009) for other training (Table 2).

Concerning patient satisfaction at the end of the consultation, this trend was confirmed with an increase of 0.22 points out of 10 points for peer training, 0.18 for association training (p=0.495 and p=0.574, respectively), and

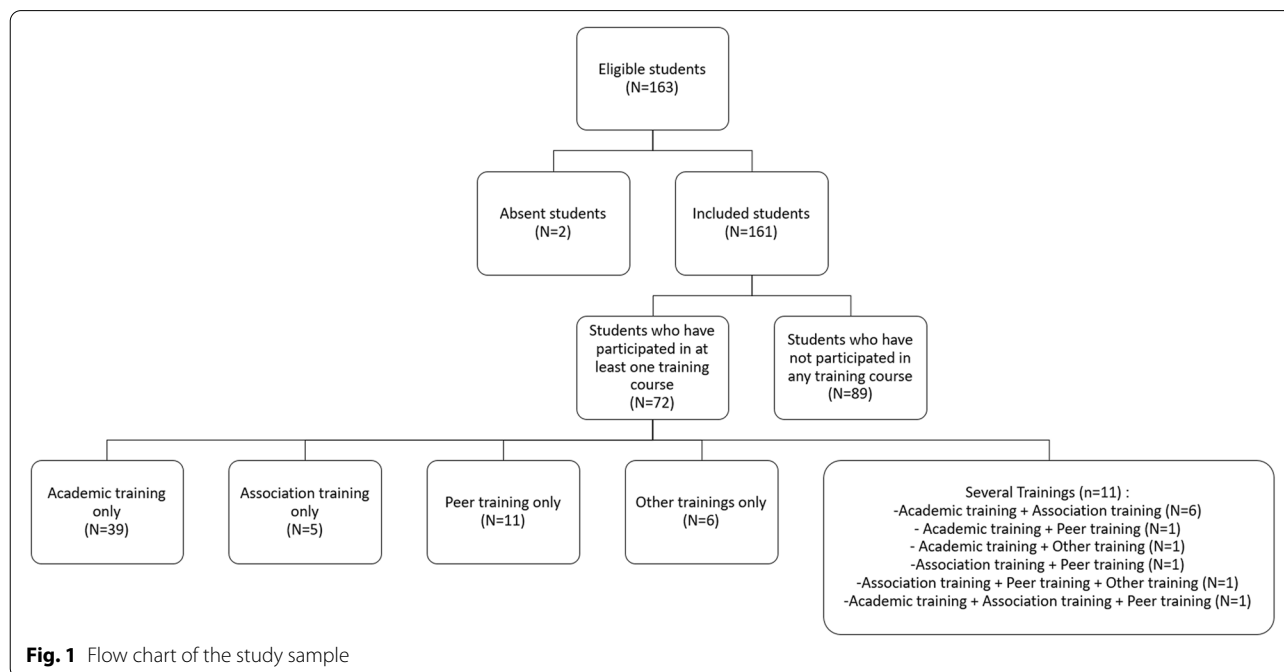


Fig. 1 Flow chart of the study sample

Table 1 Evaluation of the impact of the different training courses on the 4HCS (Four Habits Coding Scheme) and the AVS (analog visual scale)

Training	Outcomes	Intervention group Mean (SD)	Control group Mean (SD)	Unadjusted p-value	Adjusted p-value
All training combined (N=72)	4HCS (/115)	88.7 (14.3)	82.0 (18.7)	0.060	0.001
	AVS (/10)	7.91 (1.54)	7.84 (1.33)	0.661	0.779
Academic Training (N=48)	4HCS (/115)	88.6 (14.3)	83.5 (18.1)	0.080	0.002
	AVS (/10)	7.83 (1.53)	7.89 (1.39)	0.821	0.719
Association Training (N=14)	4HCS (/115)	87.5 (17.7)	84.8 (17.2)	0.573	0.296
	AVS (/10)	8.04 (1.71)	7.86 (1.40)	0.656	0.574
Peer Training (N=15)	4HCS (/115)	90.4 (15.3)	84.5 (17.3)	0.218	0.031
	AVS (/10)	8.07 (1.87)	7.85 (1.39)	0.587	0.495
Other training(N=8)	4HCS (/115)	93.2 (13.0)	84.3 (17.4)	0.075	0.007
	AVS (/10)	8.04 (1.84)	7.86 (1.39)	0.664	0.045

Table 2 Evaluation of the impact of the different training courses on the components of the 4HCS (Four Habits Coding Scheme): getting involved from the beginning (Habit 1), getting the patient's point of view (Habit 2), showing empathy (Habit 3), and being involved until the end (Habit 4)

Training	Assessment	Training (Mean and standard deviations)	Training not taken (Mean and standard deviations)	Unadjusted p-value	Adjusted p-value
All training combined (N=72)	Habit 1 (/30)	22.0 (3.51)	21.0 (4.51)	0.171	0.016
	Habit 2 (/15)	12.3 (2.38)	11.2 (2.91)	0.020	<0.001
	Habit 3 (/20)	15.0 (3.78)	14.0 (4.25)	0.146	0.002
	Habit 4 (/50)	39.0 (6.72)	36.7 (8.69)	0.078	0.004
Academic Training (N=48)	Habit 1 (/30)	21.9 (3.40)	21.2 (4.47)	0.325	0.047
	Habit 2 (/15)	12.4 (2.29)	11.3 (2.89)	0.021	0.002
	Habit 3 (/20)	15.2 (3.75)	14.0 (4.20)	0.091	0.004
	Habit 4 (/50)	39.1 (6.94)	36.9 (8.44)	0.117	0.005
Association Training (N=14)	Habit 1 (/30)	21.6 (4.34)	21.4 (4.18)	0.819	0.830
	Habit 2 (/15)	12.4 (2.47)	11.6 (2.79)	0.302	0.151
	Habit 3 (/20)	15.2 (3.47)	14.3 (4.15)	0.433	0.252
	Habit 4 (/50)	38.3 (8.84)	37.5 (8.02)	0.739	0.303
Peer Training (N=15)	Habit 1 (/30)	23.3 (3.75)	21.2 (4.19)	0.089	0.025
	Habit 2 (/15)	12.2 (2.67)	11.6 (2.77)	0.407	0.060
	Habit 3 (/20)	15.2 (4.10)	14.3 (4.10)	0.433	0.115
	Habit 4 (/50)	39.8 (6.28)	37.4 (8.20)	0.289	0.059
Other Training (N=8)	Habit 1 (/30)	22.8 (3.70)	21.3 (4.21)	0.218	0.057
	Habit 2 (/15)	12.8 (1.57)	11.5 (2.82)	0.097	0.050
	Habit 3 (/20)	16.5 (3.43)	14.2 (4.10)	0.048	0.007
	Habit 4 (/50)	41.0 (7.20)	37.3 (8.09)	0.113	0.009

0.18 points for other training ($p=0.045$) (Table 1). The mean score for all students was 7.87 (SD = 1.43).

Discussion

The training in the physician–patient relationship improved the level of interpersonal skills as evaluated by the 4-HCS score: academic training, training given by peers, and training given in a private setting on the

student's own initiative. In addition, these short courses had a significant cumulative effect: the more courses taken by the students, the more their interpersonal skills improved.

Although the results were positive, we noted that the results differed according to the type of training taken. The interpersonal skills developed were not always the same and of the same intensity, due to the

multidimensional aspect of these skills. Interpersonal skills encompass respectful attitude, attention paid to the patient, being personally present in the moment with the patient, interest in patient values and concerns, and real-time adjustment of the relationship [23]. It therefore seemed relevant to examine the mechanisms behind these differences in improvement, such as the content of the training courses and their duration. A preferred approach would be that they act not only on the communication factor in the physician–patient relationship, but also on the structuring of the interview and the empathy expressed by the health professionals. Depending on the aspect of interpersonal skills addressed, it would thus be possible to observe a difference in the increase in the specific scores of each Habit. This progression was observed more than 1 year after the training sessions, and depending on the training sessions carried out, the improvement could be felt by the patients during the simulation exercises. The evaluation of immediate satisfaction showed encouraging results, despite the fact that students have little clinical experience in the fourth year of medical studies. Students' satisfaction with the exercise could explain their full participation in the questionnaires. This questionnaire was short and had to be filled out in the continuity of the simulation, which allowed us to have no missing data.

Firstly, we were pleased with the positive results on the interpersonal skills for all the training courses. Evidence is accumulating that interpersonal skills can be acquired and improved through teaching and practice-based training for medical students [24, 25], whereas several studies have failed to demonstrate the effectiveness of training courses on these skills [26–29]. It may be interesting to ask why and how training delivered at the beginning of the medical curriculum works so well, whereas the effect obtained in continuing education is generally only moderately satisfactory [30].

Indeed, the question of temporality is also important, since the impact of training tends to diminish over time [26, 31, 32]. However, and despite the very short training courses carried out at least 1 year before our study, only the training offered by an association did not show any improvement in the interpersonal skills. However, this training took place more than 2 years before the simulated consultations (i.e., during the pre-clinical phase of the medical curriculum), which may explain its more subtle effects. The effect tested was therefore a long-term improvement effect. Internal medicine clerkship performed by medical students — between the training and the simulated consultation — include other training components than clinical reasoning such as the opportunity to acquire experience in communication with real

patients and to strengthen previously acquired interpersonal skills.

A limitation of this study was that we analyzed the practices of small groups of students in the subgroup analyses. Moreover the voluntary participation of the students in the trainings could be a bias, increasing the interest regarding the communication, the level of interpersonal skills expected is higher. The majority of students were volunteers to participate in the trainings. The lack of trainers did not allow to include all student volunteers in the different trainings.

The originality of this controlled study stems in particular from the heterogeneity of the training courses evaluated: the forms differed between courses provided by the medical faculty, courses offered by peers, or courses taken by professionals on their own initiative. The originality of this study is reinforced by the fact that the evaluation of the training impact on interpersonal skills was compared with pre-identified psychosocial risk factors. Finally, the evaluation of interpersonal skills was carried out during simulated consultations in order to be as close as possible to real-life practice conditions.

Concerning the impact of these training courses, we observed differences between the programs studied. Some of the training favored content focusing on ethics and the deontological framework and others focused on emotional experience, while some were integrated into the university curriculum and others were extra-curricular. However, the importance of the content of these programs and the environment seems to diminish with repetition. Thus, the best results were seen with the greatest number of training courses taken, even if the courses are heterogeneous. Thus, it seems that the number of training courses attended, and the regularity of these courses is more important than having sessions with standardized content. Whatever the combination of training, the effect on interpersonal skills was increased.

Research into the factors facilitating the integration of interpersonal skills in clinical practice is indispensable, as is the development of these training courses on interpersonal skills. In addition to reducing medical errors and enhancing patient safety, a beneficial effect of communication education is to increase students' self-confidence over the long term, which contributes to maintaining good mental health, especially for individuals working in stressful environments, such as medical students [15, 33, 34]. A recent meta-analysis reported a low level of evidence for studies comparing a communication skills intervention to a control group for improving medical students' interpersonal skills on only 15 studies [35]. The majority of the included studies had small group workshops (less than 70 participants) and had little to no effect on skills [35]. Effects on longer-term behaviour

change are rarely studied and our study reports some evidence. Studies with a rigorous design seem to be needed to evaluate the long-term effects of different educational approaches in a standardized way.

Conclusions

Using simulation exercises of medical consultations and a standardized evaluation of the interpersonal skills of medical students, we have shown the value of integrating training in the caregiver–patient relationship into the medical curriculum. Indeed, the impact of these training courses on students' interpersonal skills is significant, even after more than 1 year, even if they are given early in the curriculum, and even if their content is heterogeneous. The training courses are all the more effective if they are repeated. This study has potential implications for optimizing interpersonal skills training and assessment as part of the medical curriculum.

Abbreviations

4-HCS: Four Habits Coding Scheme; SD: Standard Deviation.

Acknowledgments

The authors are grateful to the staff of the Department of Performing Arts (Grenoble Alpes University) and especially to Pauline Bouchet.

Authors' contributions

AB conducted the study design and study supervision. LB, PC and AB performed data acquisition. Analysis, interpretation of data, statistical analysis and drafting the manuscript were carried out by LB and AB. Critical revision of the manuscript for important intellectual content was performed by PC. All authors read and approved the final manuscript.

Funding

None.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. The dataset used is not publicly available, as it contains links to video-recorded consultations that can identify the students involved in the study.

Declarations

Ethics approval and consent to participate

Ethics committee approval was granted on 22th February 2021, by the regional ethics committee: *Comité d'Ethique du Centre d'Investigation Clinique de Clermont-Ferrand* (IRB 00005891). All procedures performed in this study involving human participants were in accordance with the ethical standards of France and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Enrolled students provided written informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Family Medicine, University of Grenoble Alpes, Grenoble, France. ²School of Medicine, University of Grenoble Alpes, Grenoble, France. ³Clinical Epidemiology Unit, Grenoble Alpes University Hospital, CS10217,

Cedex 09, 38043 Grenoble, France. ⁴Computational and Mathematical Biology Team, TIMC UMR 5525, CNRS, University of Grenoble Alpes, Grenoble, France.

Received: 12 May 2021 Accepted: 27 January 2022

Published online: 22 February 2022

References

- Krupat E, Frankel R, Stein T, Irish J. The four habits coding scheme: validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns*. 2006;62:38–45.
- Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J. Cross-cultural adaptation of the 4-habits coding scheme into French to assess physician communication skills. *PLoS One*. 2020;15(4):e0230672.
- Candib LM. When the doctor takes off her clothes: reflections on being seen by patients in the exercise setting. *Fam Syst Health*. 1999;17(3):349–63.
- Aper L, Reniers J, Koole S, Valcke M, Dereze A. Impact of three alternative consultation training formats on self-efficacy and consultation skills of medical students. *Med Teach*. 2012;34(7):e500–7.
- Bellier A. Associated factors with interpersonal skills in initial training during standardized medical consultations: Faculté de Médecine de Grenoble: Université Grenoble Alpes; 2019.
- Libbrecht N, Lievens F, Carette B, Côté S. Emotional intelligence predicts success in medical school. *Emot Wash DC*. 2014;14(1):64–73.
- Kohn LT, Corrigan JM, Donaldson MS, editors. Institute of Medicine (US) committee on quality of health Care in America. To err is human: building a safer health system. Washington (DC): National Academies Press (US); 2000.
- Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: description of the research domain. *Soc Sci Med*. 1998;47(10):1573–88.
- Sanson-Fisher R, Hobden B, Waller A, Dodd N, Boyd L. Methodological quality of teaching communication skills to undergraduate medical students: a mapping review. *BMC Med Educ*. 2018;18(1):151.
- Puntis M. Skills for communicating with patients (2nd edn). *Ann R Coll Surg Engl*. 2006;88(1):85–6.
- Rao JK, Anderson LA, Inui TS, Frankel RM. Communication interventions make a difference in conversations between physicians and patients: a systematic review of the evidence. *Med Care*. 2007;45(4):340.
- Gulbrandsen P, Jensen BF, Finset A, Blanch-Hartigan D. Long-term effect of communication training on the relationship between physicians' self-efficacy and performance. *Patient Educ Couns*. 2013;91(2):180–5.
- Batt-Rawden SA, Chisolm MS, Anton B, Flickinger TE. Teaching empathy to medical students: an updated, systematic review. *Acad Med J Assoc Am Med Coll*. 2013;88(8):1171–7.
- Bas-Sarmiento P, Fernández-Gutiérrez M, Baena-Baños M, Corro-Bermejo A, Soler-Martins PS, de la Torre-Moyano S. Empathy training in health sciences: a systematic review. *Nurse Educ Pract*. 2020;44:102739.
- Buckley S, Ambrose L, Anderson E, Coleman JJ, Hensman M, Hirsch C, et al. Tools for structured team communication in pre-registration health professions education: a best evidence medical education (BEME) review: BEME guide no. 41. *Med Teach*. 2016;38(10):966–80.
- Stein T, Frankel RM, Krupat E. Enhancing clinician communication skills in a large healthcare organization: a longitudinal case study. *Patient Educ Couns*. 2005;58(1):4–12.
- Jensen Bard F, Pal G, Dahl FA, Krupat E, Frankel RM, Finset A. Effectiveness of a short course in clinical communication skills for hospital doctors: results of a crossover randomized controlled trial (ISRCTN22153332). *Patient Educ Couns*. 2011;84(2):163–9.
- Fossli Jensen B, Gulbrandsen P, Benth JS, Dahl FA, Krupat E, Finset A. Interrater reliability for the four habits coding scheme as part of a randomized controlled trial. *Patient Educ Couns*. 2010;80(3):405–9.
- Gulbrandsen P, Krupat E, Benth JS, Garratt A, Safran DG, Finset A, et al. "Four habits" goes abroad: report from a pilot study in Norway. *Patient Educ Couns*. 2008;72(3):388–93.
- Scholl I, Nicolai J, Pahlke S, Kriston L, Krupat E, Härter M. The German version of the four habits coding scheme – association between physicians' communication and shared decision making skills in the medical encounter. *Patient Educ Couns*. 2014;94(2):224–9.

21. Gedda M. Traduction française des lignes directrices STROBE pour l'écriture et la lecture des études observationnelles. *Kinésithérapie Rev.* 2015;15(157):34–8.
22. Asan O, Montague E. Using video-based observation research methods in primary care health encounters to evaluate complex interactions. *Inform Prim Care.* 2014;21(4):161–70.
23. Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, et al. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. *Acad Med J Assoc Am Med Coll.* 2004;79(6):495–507.
24. Yedidia MJ, Gillespie CC, Kachur E, Schwartz MD, Ockene J, Chepaitis AE, et al. Effect of communications training on medical student performance. *JAMA.* 2003;290(9):1157–65.
25. Smith S, Hanson JL, Tewksbury LR, Christy C, Talib NJ, Harris MA, et al. Teaching patient communication skills to medical students: a review of randomized controlled trials. *Eval Health Prof.* 2007;30(1):3–21.
26. Junod Perron N, Sommer J, Louis-Simonet M, Nendaz M. Teaching communication skills: beyond wishful thinking. *Swiss Med Wkly.* 2015;145:w14064.
27. Liénard A, Merckaert I, Libert Y, Bragard I, Delvaux N, Etienne A-M, et al. Transfer of communication skills to the workplace during clinical rounds: impact of a program for residents. *PLoS One.* 2010;5(8):e12426.
28. Heaven C, Clegg J, Maguire P. Transfer of communication skills training from workshop to workplace: the impact of clinical supervision. *Patient Educ Couns.* 2006;60(3):313–25.
29. van den Eertwegh V, van Dulmen S, van Dalen J, Scherpbier AJJA, van der Vleuten CPM. Learning in context: identifying gaps in research on the transfer of medical communication skills to the clinical workplace. *Patient Educ Couns.* 2013;90(2):184–92.
30. Rotthoff T, Baehring T, David DM, Bartrick C, Linde F, Willers R, et al. The value of training in communication skills for continuing medical education. *Patient Educ Couns.* 2011;84(2):170–5.
31. Aspegren K. BEME guide no. 2: teaching and learning communication skills in medicine—a review with quality grading of articles. *Med Teach.* 1999;21(6):563–70.
32. van Dalen J, Kerkhofs E, van Knippenberg-Van Den Berg BW, van Den Hout HA, Scherpbier AJJA, van der Vleuten CPM. Longitudinal and concentrated communication skills programmes: two dutch medical schools compared. *Adv Health Sci Educ Theory Pract.* 2002;7(1):29–40.
33. Brown JB, Boles M, Mullooly JP, Levinson W. Effect of clinician communication skills training on patient satisfaction. A randomized, controlled trial. *Ann Intern Med.* 1999;131(11):822–9.
34. Omura M, Maguire J, Levett-Jones T, Stone TE. Effectiveness of assertive communication training programs for health professionals and students: a systematic review protocol. *JBI Database System Rev Implement Rep.* 2016;14(10):64–71.
35. Gilligan C, Powell M, Lynagh MC, Ward BM, Lonsdale C, Harvey P, et al. Interventions for improving medical students' interpersonal communication in medical consultations. *Cochrane Database Syst Rev.* 2021;2:CD012418.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



4.2. **Partie 8)** Protocole d'un essai contrôlé randomisé pour l'évaluation d'un programme pédagogique dédié aux compétences relationnelles en formation continue

Article publié — Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): Protocol for a randomised controlled trial

Dans cet article publié le 15 février 2022 dans la revue *BMJ Open*, nous avons développé un protocole d'évaluation de l'efficacité d'une formation pédagogique multifacette dédiée aux compétences relationnelles, pour les médecins hospitaliers. Cette formation a été créée pour les besoins de cet essai contrôlé randomisé en s'appuyant sur l'échelle 4-HCS et les programmes de formation montés autour de celle-ci. Par ailleurs, des notions de Process-Communication ont été introduites avec un retour individuel sur des consultations réelles.

Cette étude a été financée dans le cadre de l'appel d'offres interne du CHU Grenoble Alpes.

Introduction

Les compétences relationnelles, qui englobent la communication et l'empathie, sont des éléments clés d'une consultation médicale efficace. Bien que de nombreuses organisations aient mis en place des programmes de formation structurés, il existe peu de preuves de leur efficacité à améliorer les compétences relationnelles des médecins. Cette étude vise à évaluer l'efficacité d'un programme standardisé multifacette de développement des compétences relationnelles, destiné aux médecins hospitaliers.

Méthodes et analyse

Cette étude est un essai de supériorité prospectif, randomisé (avec un ratio d'allocation 1:1), contrôlé, en ouvert, à deux bras parallèles, mené dans un seul hôpital universitaire. Les médecins seront randomisés pour recevoir soit un programme de formation à facettes multiples, soit aucune intervention. L'intervention combine deux sessions de formation de 4 heures, la diffusion de matériel éducatif interactif, l'évaluation de consultations vidéo-enregistrées et un feedback individuel. Le critère d'évaluation principal est le score global de l'échelle 4-HCS, évalué par deux évaluateurs indépendants, en aveugle du bras de l'étude, sur la base de consultations vidéo-enregistrées, avant et après l'intervention. Les critères de jugement

secondaires comprennent la satisfaction du patient, l'alliance thérapeutique, le développement personnel du médecin au travail et la durée de la consultation médicale.



Éthique et diffusion

Le protocole de l'étude a été approuvé le 21 octobre 2020 par le CECIC Rhône-Alpes Auvergne, Clermont-Ferrand, France (IRB 5891). Tous les participants ont fourni un consentement éclairé écrit. Des efforts seront faits pour publier les résultats primaires dans les 6 à 9 mois suivant la fin de l'étude, qu'ils confirment ou infirment l'hypothèse de recherche.

Numéro d'enregistrement de l'essai

NCT04703816.

BMJ Open Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): protocol for a randomised controlled trial

Alexandre Bellier ^{1,2,3}, José Labarère ^{1,2,4}, Zaza Putkaradze,³ Guillaume Cavalié,⁵ Sylvain Carras,⁵ Félix Pelen,⁵ Adeline Paris,³ Philippe Chaffanjon⁴

To cite: Bellier A, Labarère J, Putkaradze Z, *et al*. Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): protocol for a randomised controlled trial. *BMJ Open* 2022;**12**:e051600. doi:10.1136/bmjopen-2021-051600

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-051600>).

Received 23 March 2021
Accepted 04 January 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Alexandre Bellier;
abellier@chu-grenoble.fr

ABSTRACT

Introduction Interpersonal skills, encompassing communication and empathy, are key components of effective medical consultations. Although many organisations have implemented structured training programmes, limited evidence exists on their effectiveness in improving physician interpersonal skills. This study aims to evaluate the effectiveness of a standardised, multifaceted, interpersonal skills development programme for hospital physicians.

Methods and analysis This study is a prospective, randomised (with a 1:1 allocation ratio), controlled, open-label, two parallel arm, superiority trial conducted at a single university hospital. Physicians will be randomised to receive either a multifaceted training programme or no intervention. The experimental intervention combines two 4-hour training sessions, dissemination of interactive educational materials, review of video-recorded consultations and individual feedback. The primary outcome measure is the overall 4-Habits Coding Scheme score assessed by two independent raters blinded to the study arm, based on video-recorded consultations, before and after intervention. The secondary outcomes include patient satisfaction, therapeutic alliance, physician self-actualisation and the length of medical consultation.

Ethics and dissemination The study protocol was approved on 21 October 2020 by the CECIC Rhône-Alpes Auvergne, Clermont-Ferrand, France (IRB 5891). All participants will provide written informed consent. Efforts will be made to release the primary results within 6 to 9 months of study completion, regardless of whether they confirm or deny the research hypothesis.

Trial registration number NCT04703816.

INTRODUCTION

Background

The doctor–patient relationship is central to medical practice and its quality can have a direct impact on patient outcomes.¹ The quality of the interaction between physician and patient during a consultation is a

Strengths and limitations of this study

- Physician's interpersonal skills is a major determinant of patient satisfaction with medical consultation and compliance with plan of care.
- The impact of interpersonal skill training will be studied from both the patient's and the physician's perspective.
- Our study is designed as a randomised controlled trial in order to provide the highest level of evidence on the effectiveness of interpersonal skill training programme.
- Participating physicians cannot be blinded to study intervention in this open-label trial.
- Video recording of medical consultations may hamper physician and patient participation in the trial.

major determinant of patient satisfaction and adherence to the plan of care. Interpersonal skills, such as patient-centred communication and empathy, are of considerable importance in establishing the unique relationship between doctor and patient, at a time when medical practice is increasingly focused on the technical act of care. Communication is recognised as an essential skill for effective medicine.^{2–4} Interpersonal skills are defined by the presence of effective verbal and non-verbal behaviours in the context of individual interactions with patients or families.⁵

However, a decline in communication skills among physicians over the course of their careers⁶ and a decline in empathy⁷ have been reported, despite the importance of these non-technical skills.

Interpersonal skill training program

Many organisations have implemented training programmes and routinely assess

physicians' communication skills using standardised scales.^{2,8} However, limited evidence exists on the effectiveness of these programmes in improving physician interpersonal skills. Indeed, the vast majority of published reports are descriptive in design, lack adequate control groups, enrolled medical students or had methodological weaknesses.^{9,10} Less than 2% of published studies are randomised controlled trials¹⁰ and the best strategy for improving physician interpersonal skills remains to be determined.⁶

Evidence is currently lacking on the effectiveness of training programme in altering patient outcomes.¹¹ Few studies have shown an impact of improved physician interpersonal skills on patient satisfaction^{12,13} and even fewer investigated the effect on therapeutic alliance, which is correlated with the quality of doctor–patient communication.¹⁴

The 'Four Habits Model' is a training programme addressing basic medical interview tasks that was developed within the US Kaiser Permanente Health Maintenance Organization. This training programme has been implemented for teaching effective communication skills in various organisations in the USA and Norway.¹² Previous reports suggest that training programmes based on the Four Habits Model may improve physicians' communication self-efficacy in the long term¹⁵ and patient satisfaction with medical consultation.¹²

Finally, physician interpersonal skills might be improved at the price of longer medical consultations. Substantial heterogeneity exists in the length of medical consultation across countries, ranging from less than 10 min in the UK to more than 20 min in the USA, with an intermediate value of 16 min in France.¹⁶ Longer medical consultations generate extra costs and the length of consultation has been shown to relate to the economic expenditure per capita of the country.¹⁶ Yet, it remains uncertain whether the length of consultation is associated with physician performance and patient satisfaction.¹⁷

Research hypothesis

The primary hypothesis guiding the project is that a multifaceted structured training programme may improve the communication and interpersonal skills of hospital physicians, without altering the length of consultation. A multifaceted programme combines two or more components. Although speculative, multifaceted interventions may be more effective than single-component interventions in changing physician interpersonal skills. Our experimental multifaceted intervention will combine learning techniques for continuing medical education, role plays for practice and feedback on individual performance. Our secondary hypotheses are that improved physician interpersonal skills are paralleled by (1) increased levels of patient satisfaction with medical consultation and therapeutic alliance and (2) changes in physician professional fulfilment and self-actualisation.

Objectives

We propose to conduct an experimental study with the highest level of scientific evidence (randomised controlled trial) to determine whether a multifaceted training programme improves physician interpersonal skills with a positive impact on patient outcomes. The Four Habits Model forms the framework of the experimental intervention.^{12,18} This multifaceted intervention will combine theoretical and practical training sessions with the use of video-recorded medical consultations and personalised feedback on individual performance during medical consultations.

The primary objective of the study is to determine whether a multifaceted training programme is effective in improving physician interpersonal skills as rated with the 4-Habits Coding Scheme (HCS) relative to baseline measure in comparison with a control group receiving no intervention.

The secondary objectives of the study are to compare patient satisfaction, patient therapeutic alliance, physician personal achievement and the length of consultation between the experimental and control groups.

METHODS

Trial design

To ensure a high level of evidence, we designed a prospective superiority randomised controlled intervention trial. To prevent unintentional spill-over of intervention effect from experimental to control arm, the unit of randomisation will be physician. Given the educational nature of the intervention, physicians cannot be blinded to the study group; however, the patients, the raters in charge of coding the 4-HCS based on video-recorded consultations and the statistician in charge of the primary and secondary outcome analysis will be blinded to study group.

Study settings

The project is conducted at a single university-affiliated public acute care hospital in France.

Recruitment of clinicians

Each physician board-certified in medical, surgical or gynaecology-obstetrics specialty at Grenoble Alpes University Hospital was invited to participate in the study. Physicians were contacted by electronic mails send by the principal investigator (AB). Contact information was retrieved from the hospital database of professional electronic addresses. Correspondence enclosed a cover and the study protocol. A reminder was e-mailed to non-respondents 1 month later. Posters calling for volunteers were also displayed in areas frequented by physicians in the hospital. The principal investigator has no power relationship with the physicians participating in the study. Of 839 physicians contacted by electronic mail, 37 volunteered to participate, and 28 were recruited.

Physicians volunteering to participate are required to meet the inclusion and exclusion criteria. Prior to

enrolment, all participating physicians will be asked to provide written informed consent.

Patient recruitment

Consecutive adult outpatients will be screened for eligibility if they consult with a physician participating in the study. To be eligible, patients will be required to meet all four inclusion criteria and none of the exclusion criteria. Participating physician will be required to recruit eight consecutive eligible patients from their scheduled consultations. The recruitment period will extend to the physician's inclusion of four patients in the preintervention period and four patients in the postintervention period, respectively. If the physician leaves the study before the intervention is implemented, he or she will be excluded from the study. If the physician leaves the study after the intervention is implemented, the data acquired so far will be retained unless the physician objects.

In order to quantify the likelihood of possible bias in patient selection, a list of consultations during the recruitment period will be established for each participating physician. This list will include the patient's age and gender as well as the reason for exclusion.

The study was planned to include patients from 1 July 2021 to 31 October 2021, with an estimated trial end date of 31 December 2021.

Eligibility criteria

Inclusion criteria

- ▶ Physicians:
 - Physicians board certified in medical, surgical or gynaecology-obstetrics specialty at Grenoble Alpes University Hospital.
 - Provision of written informed consent.
- ▶ Patients:
 - Scheduled consultation in the public sector at Grenoble Alpes University Hospital.
 - Patient treated in the participating physician's department.
 - Initial consultation for new patient.
 - Age ≥ 18 years old.

Exclusion criteria

- ▶ Physicians:
 - Problems expressing or understanding the French language for cultural or language reasons.
- ▶ Patients:
 - Patient with difficulties in understanding, expressing or reading the French language for cultural or language reasons.
 - Patient who is unable to provide written informed consent, because of cognitive impairment, altered mental status or communication impairments for medical reason.
 - Patient subject to a legal protection measure or unable to express their objection.

The potential for recruiting physicians into this study was assessed beforehand by interviewing physicians who

participated to the activities of the continuing medical education department at Grenoble University Hospital.

Interventions

Inclusion visit

During the inclusion visit, the volunteer physician is asked to meet with one of the study investigators to obtain consent and to report his or her specialty (medicine, surgery or gynaecology-obstetrics) and status (incumbent or non-incumbent).

Prior to the consultation, eligible patients are contacted by phone to be informed about the study protocol and their potential participation. At the time of the medical consultation, the patient receives additional information about the study by a research team member. A generic notice on internal data search is given to the patient. The research team member checks for the absence of any objection. Patient demographics and medical baseline characteristics are collected using a self-administered questionnaire.

Preintervention study period

Video-recording equipment will be provided to participating physicians. The physician will start the video recording using a miniaturised recording device placed on the desk, before picking up the patient in the waiting room, by simply pressing the recording button. The physician will end the recording in the same way at the end of the medical consultation. The video recording will, therefore, be centred on the desk making the doctor and the patient visible, with the notable exception of the clinical examination table.

Practitioners are invited to videotape four medical consultations with consecutive eligible outpatients over a 3-month period. After consultations, satisfaction and therapeutic alliance, self-administered questionnaires will be given to the participating patient with a stamped return envelope. A reminder will be made by phone to non-respondents within 15 days of consultation. Questionnaires sent back within 30 days of medical consultations will be included in the analysis. The participating physician will be invited by mail to fill in the personal achievement questionnaire.

Experimental training programme

The physicians assigned in the intervention arm will receive the experimental multifaceted training programme. Physicians assigned in the control group will not receive any specific intervention. The theoretical model of the intervention is based on Philip Price's benchmark of the attributes of being a good practicing physician¹⁹ and on the skills associated with the patient-centred relationship.²⁰ Each of the dimensions of the 4-HCS (ie, 'Invest in the beginning,' 'Elicit Patient's Perspective,' 'Demonstrate empathy,' 'Invest in the end') is the subject of specific work during the workshops. For the conceptual framework of the intervention, we will focus on training in interpersonal skills, including

communication and ethics based on the extensive experience of Kaiser Permanente and the Bayer Institute for Healthcare Communication.^{12 18} The overall effectiveness of the programme has undergone preliminary evaluations but no analysis on a component-by-component has been performed.^{13 15} We have adopted the Cochrane Effective Practice and Organisation of Care (EPOC) group typology to present our programme. In detail, the intervention consists of training by an expert in the field of communication and interpersonal skills with experience in the hospital medical field. This expert will be accompanied by a physician with experience in the evaluation of interpersonal skills for coanimation. The training will comprise two 4-hour sessions with a 1-month interval in between. Prior to the first workshop, a questionnaire will be sent to each doctor to identify the profile of the practices of the different professionals and to adapt the discourse and the workshops. The first 4-hour session of training will, thus, include a review of the skills needed to establish a patient-centred relationship, using, in particular, the various essential points assessed by the 4-HCS scale.²¹ An introduction to active listening and Process Communication techniques will also be provided with the dissemination of educational and interactive materials. The Process Communication model developed by the psychologist Taibi Kahler makes it possible to identify one's own communication profile and that of the patient in order to adapt communication. The workshop provides an understanding of how to enter into a relationship, how to analyse non-verbal behaviour and how to improve patient-centred communication. Then, the second half-day of training will consist of working on interpersonal skills in relation to the communication techniques developed in the first workshop, putting them into practice through role playing. Finally, difficult, emotionally charged consultations and reactions under stress will be addressed, with specific techniques for dealing with them. These different workshops are inspired by Kaiser Permanente's experience of more than 20 years in the USA¹² and by Norwegian hospital teams.¹⁵ Participating physicians will then receive individual feedback on their interpersonal skills analysed via the 4-HCS scale²¹ on the basis of video-recorded consultations. The complete description of the educational programme is described in [table 1](#) according to the template for intervention description and replication checklist.²² This description follows the taxonomy for delivery characteristics proposed by Schulz *et al.*²³

Postintervention study period

At the end of the second workshop, physicians assigned in the intervention arm will be provided with personalised feedback on the acquisition of interpersonal.

Physicians assigned in the control group will not receive any specific training or feedback during the postintervention study period. Patients enrolled by physicians assigned in the control group will receive usual care. Physicians assigned in the control arm will not be exposed to any

component of the multifaceted intervention during the conduct of the study, in order to minimise the likelihood of unintentional contamination from experimental to control group, in this parallel-arm cluster randomised trial. The participating physicians in the two study arms will be invited to videotape medical consultations with at least four consecutive eligible patients over a 3-month period.

At the end-of-study visit, one of the study investigators who assessed the interpersonal skills will provide personalised feedback to each participating physician and will note any changes in interpersonal skills during the consultations, for the intervention and control arms.

The physicians assigned in the control arm will benefit from the experimental intervention at the end of the trial, if they wish.

Outcomes

Primary outcome measure

The primary outcome measure is the overall score produced by the cross-cultural adaptation of the 4-HCS scale in French.²¹ The 4-HCS was cross-culturally adapted by conducting forward and backward translations with independent translators from the original scale,²⁴ following international guidelines.²⁵ Cronbach's alpha was 0.94 for the overall 4-HCS, ranging from 0.72 to 0.88 across subscales. Median average absolute-agreement intraclass correlation coefficient estimates were 0.74 (range, 0.68–0.84) and 0.85 (range, 0.76–0.91) for inter-rater and intrarater reliability of habit subscales, respectively.²¹

Two independent raters blinded to study arm assessed physician interpersonal skills based on video-recorded consultations. The raters will be the same as those involved in the cross-cultural adaptation of the 4-HCS in French,²¹ to ensure a satisfactory level of reliability. The experts will receive all the videos for the period concerned at random. A random list of videos will be produced by experts for the first study period and then for the second period to allow individual feedback on the interpersonal skills of the physicians in the intervention group (at the end of the first and second periods). Each video-recorded consultation will be analysed within 30 days of acquisition.

Secondary outcome measure

The secondary patient-level outcome measures include patient satisfaction, therapeutic alliance and the length of consultation. Patient satisfaction with the medical consultation will be assessed with the cross-cultural adaptation of the American Board of Internal Medicine Patient Satisfaction Rating Scale in French.²⁶ Patient therapeutic alliance will be measured using the cross-cultural adaptation of the Inventory of the Therapeutic Alliance in French.²⁷ The optimal recall period for measuring patient satisfaction with medical consultation is controversial. The criteria that guided our choice of recall period (up to 30 days after the consultation) were (1) patient ability to easily and accurately recall the information requested at home, (2) the

Table 1 Intervention description according to the TIDieR checklist (template for intervention description and replication)

Brief name	Multifaceted programme for interpersonal and communication skills development in medical consultation
Why	Improved doctor–patient interpersonal skills are associated with improved patient satisfaction and quality of care, but there is a lack of evidence in the literature on how to develop these skills.
What	<p>The multifaceted programme includes two 4-hour workshops and feedback on the interpersonal skills observed during the doctor's consultation. Before the first workshop, an evaluation questionnaire based on the Process Communication model is sent to each participant. This questionnaire allows us to establish the communication profile of each participant. The first workshop presents the Process Communication theoretical model of communication during 2 hours to explain the profile of each person. A 1-hour theoretical presentation is also given on interpersonal skills, based on the 4-HCS scale and the model developed by Kaiser Permanente organisation. The last hour consists of a communication approach based on Process-Com and adapted to the doctor–patient relationship, linking the two theoretical models presented.</p> <p>The second workshop includes role-playing situations in groups of three people, with an observer, a physician and a patient. An observation grid inspired by the 4-HCS scale is given to each observer to allow a constructive debriefing on interpersonal skills. The participants take turns exchanging roles and a collective debriefing is conducted after each clinical situation. These clinical situations involve different communication profiles in order to apply the knowledge acquired in the first workshop.</p> <p>A detailed written analysis of the interpersonal skills observed during the consultations is finally given to each participant after the workshops. This analysis details strengths and areas for improvement, based on the 4-HCS assessment of the video recorded consultations by the physicians.</p>
Who provided	The workshops are conducted by an expert in the field of communication with 20 years of experience in the hospital medical field. This expert is a professional trainer with a degree in communication and expert in the Process Communication model. The physician who also conducts the training is a physician who has conducted the cross-cultural adaptation of the 4-HCS scale into French, with experience in nearly 1000 consultation assessments using this scale. Interpersonal skills assessments are conducted by another physician with experience of several hundred evaluated consultations with 4-HCS scale.
How	The workshops are conducted in groups of 8–12 people with 2 trainers at 1 month intervals. The evaluations of the participants' consultations are sent by e-mail in the form of paragraphs describing the strengths and weaknesses in relation to the interpersonal skills assessed by the 4-HCS scale. Videos are added to the e-mail.
Where	The workshops take place in a classroom located in the hospital. Medical consultations take place in the doctor's usual department.
When and how much	The training includes two workshops of 4 hours at 1 month interval, as well as individual feedback on 8 consultations of the participating physician.
Tailoring	The training is adapted to the communication profile of each participant during the first workshop, based on the results of the previously completed Process Communication questionnaires. The feedback during the second workshop is adapted to the content observed during the different role plays.
Modifications	No changes made to the programme
How well (planned)	The verification that each workshop participant has completed the communication profile questionnaire is done prior to the training. A monitoring is also done during the second workshop by the trainers to ensure that each participant changes roles systematically during the role-playing session.

HCS, Habits Coding Scheme; TIDieR, Template for Intervention Description and Replication.

potential for maturation bias and (3) the consistency with previous studies.¹⁸ The length of medical consultation will be quantified by the two independent raters based on the video recording. The physician-level secondary outcome measures include the subscale score for each of the four dimensions of the cross-cultural adaptation of the 4-HCS in French and self-actualisation assessed using the French-language cross-cultural adaptation of the Maslach Burnout Inventory multidimensional scale.²⁸

Sample size

A sample of 56 patients included by 14 physicians (average number of patients/physician: four patients/physician)

in each arm (ie, 112 patients/28 physicians) would confer a power greater than 80% to show an average difference of 7.5 points in the 4-HCS score (two-sided alpha level of 0.05). This sample size was calculated under the hypothesis of a SD of the 4-HCS score equal to 10²⁴ and an intra-cluster correlation coefficient equal to 0.30.

Each arm of the trial will include 56 preintervention and 56 postintervention patients, for a total of 224 patients. This number makes it possible to show a significant interaction term between the trial arm and period equal to 0.30, with a power greater than 80% and an inflation factor equal to 1.9.²⁹

Recruitment

A member of research team working at the Clinical Investigation Center (Grenoble Alpes University Hospital) will recruit study participants.

Randomisation

The unit of randomisation is the physician, in order to minimise the likelihood of cross-contamination between study arms. Randomisation will be stratified and balanced by minimisation on the status (incumbent vs non-incumbent) and specialty (medical vs surgical) of the participating physicians. We are anticipating that incumbent versus non-incumbent status and specialty are baseline physician characteristics that may confound the effectiveness of the experimental intervention in improving interpersonal skills. An independent statistician will generate allocation sequence, with a 1:1 ratio using computer-generated random numbers. To ensure concealment, study arm will not be released during the preintervention period. The randomisation will be centralised at the Clinical Investigation Centre of Grenoble Alpes University Hospital. The moment of physician randomisation will take place at the end of the pre-intervention period.

Allocation and blinding

Participating physicians cannot be blinded to study intervention in this open-label trial. However, the patients, the raters evaluating video-recorded consultations and the statistician in charge of the primary and secondary outcome analyses will be blinded to the study arm. The correspondence between the anonymity number and the allocation group with the arm of the study can be determined only by the statistician who generates the sequence of randomisation. The physician will be explicitly asked not to disclose to the patient whether or not he or she is assigned to the experimental intervention.

Data collection, data management and confidentiality

An electronic case report form will be created for the study. Trial data management will be carried out in accordance with on-site standard operating procedures. A data management plan will be developed by the data manager and approved by the principal investigator, the scientific coordinator and the study statistician. Different approaches will be implemented to optimise data quality and identified in a data validation plan including routine checks (valid values, range checks and consistency checks) at the time of data entry for specific fields, double data entries, execution of computerised programmes for the detection of additional inconsistencies, follow-up at regular intervals of requests for corrections and final review of the data prior to locking the database. The collected data will be stored in areas with limited access. Confidentiality of data, including the personal data and video recording, will be maintained.

Statistical methods

A statistical analysis plan (SAP) will be developed prior to database lock, reviewed by the principal investigator and an independent statistician and approved by the steering committee. Any post hoc or unplanned analyses not specified in the SAP will be clearly identified as such in the final statistical report and manuscripts for publication. No formal interim analysis is planned.

The intention-to-treat (ITT) population will consist of all observations for participating physicians who have been randomised. Patients and physicians will be analysed in the study arm assigned by randomisation. The per-protocol (PP) population will consist of all observations for randomised physicians without any major deviation from the protocol (non-compliance with the multifaceted training programme) and evaluable. The numbers of patients and physicians in ITT and PP populations will be presented by study arm throughout a flowchart extension for cluster randomised trials.

Descriptive summary statistics will be used for reporting continuous (arithmetic mean and SD or median and 25th–75th percentiles) and categorical (numbers and percentages) variables. Baseline and demographic characteristics will be summarised for both ITT and PP populations. Baseline patient and physician characteristics will be compared between the two study arms.

The primary outcome analysis (ie, 4-HCS overall score) will be conducted within the ITT population and, for sensitivity reason, repeated within the PP population. We will use a difference-in-differences approach. To account for patient clustering within participating physicians, we will analyse 4-HCS overall score using random-intercept linear regression model for continuous dependent variable.

The analysis of secondary outcomes will be exploratory in nature. Inferential comparisons for participating physicians between study arms will be performed using the t test or Wilcoxon rank-sum test for unpaired data for continuous outcome variables. To account for patient clustering within participating physicians, we will analyse secondary outcome measures using random-intercept linear regression model for continuous dependent variable.

No subgroup analysis is planned for the primary and secondary study outcomes.

For transparency purpose, the completeness of study data will be reported for baseline characteristics and outcome variables. In cases of participating physician withdrawal, we are planning to perform multiple imputation of missing data. To assess the robustness of our findings, we will perform multivariate imputation using chained equations for imputing missing primary and secondary outcome values.³⁰

All primary and secondary outcome analyses will be performed on both ITT and PP populations at a two-sided alpha level of 0.05. All statistical analyses will be performed with Stata Special Edition V.16 or higher (Stata Corporation, College Station, Texas) and RStudio V.1.3.959 or higher (PBC, Boston, Massachusetts).

Additional software may be used for the production of graphics and for statistical methodology not provided by these software packages.

Data monitoring

Monitoring involves onsite periodic reviews of core trial processes and documentation conducted by staff appointed by the sponsor (Grenoble Alpes University Hospital). The sponsor may require an audit in order to obtain independent appraisal of trial data quality and integrity.

Patients and public involvement statement

Patient and the public representatives are not involved in the study design, recruitment, conduct or dissemination of findings.

Research checklist

The present protocol complies with the *Standard Protocol Items: Recommendations for Interventional Trials* 2013 statement.³¹

ETHICS AND DISSEMINATION

Research ethics approval

The study protocol was approved on 21 October 2020 by the CECIC Rhône-Alpes Auvergne, Clermont-Ferrand, France (IRB 5891). All participants will provide written informed consent.

Protocol amendments

During the conduct of the study, protocol changes are not desirable and should not be made unless new information strongly suggests that such changes would strengthen the scientific validity of the study. If substantive modifications are necessary that may impact on the study conduct or results, including changes of study objectives, eligibility criteria, data collection methods, variable definitions or significant administrative aspects, they will require a formal amendment to the protocol. The date, description of changes and rationale for amendments will be reported in a tabular format. Minor corrections or clarifications that have no effect on the way the study is to be conducted will be documented in a memorandum.

Protocol registration

Recorded information will be updated on a regular basis.

Consent or assent

Before participating in the trial, the patient will be informed of all pertinent aspects of the study (including objective, design, methods, constraints, anticipated risks and benefits), be provided with information form and be given time to ask questions and time to consider the decision to participate. The patient will be informed that the quality of care will not be affected by the decision to participate in or withdraw from the study. The investigator is responsible for obtaining informed consent for participating in the study and for image and voice right before any study intervention

is administered. The acquisition of informed consent will be documented in the patient's medical records, and the informed consent form will be signed and personally dated by the patient and by the investigator.

Dissemination policy

Efforts will be made to reduce the interval between data collection completion and the release of the primary study results. The results of this study will be published, regardless of whether they confirm or deny the research hypothesis. It is expected that 6–9 months will be necessary to compile the primary study results before manuscript submission to an appropriate journal. All publications will comply with the Consolidated Standards of Reporting Trials extension to cluster randomised trial guidelines, as appropriate.³² All investigators and subinvestigators that have actively participated in the trial will be listed at the end of all manuscripts if this can be arranged with the publisher. Authors' names will be listed in order of contribution. Assistance for preparing and editing manuscripts (ie, English language revision) provided by professional medical writers will be acknowledged.

No later than 3 years after final acceptance of the primary study paper, a completely deidentified data set will be available for sharing purpose, on reasonable request to the principal investigator. In accordance with French regulation, study participants will be provided with the overall trial results on request to the principal investigator.

DISCUSSION

This protocol describes the rationale for the EPECREM (Effectiveness of a multifaceted Program to Evaluate the improvement of RElational Competencies in Medical consultation) randomised controlled trial project, explains how the experimental intervention will be implemented, how data collection will be conducted and how the results will be analysed and interpreted. The potential limitations of this trial deserve mention. First, the control group will not receive any specific intervention. Actually, our trial is not designed to compare the effectiveness of concurrent training programmes but to demonstrate that a multifaceted training programme improves physician interpersonal skills. Second, physicians might avoid recruiting patients with whom the interaction is perceived as unfavourable. To limit the potential for patient selection bias, participating physicians will be invited to enrol consecutive eligible patients. Only initial consultations for new patients will be eligible. A list of eligible consultations during the recruitment period will be established for each participating physician. Third, the Maslach Burnout Inventory scale was originally developed for assessing burnout and may lack sensitivity to detect clinically significant differences in physician self-actualisation between study arms. To our knowledge, very few standardised scales assessing physician's self-actualisation have been published. The Maslach Burnout Inventory, which has been translated and validated in French, includes a self-actualisation subscale. Fourth, our study is conducted at

a single university-affiliated hospital in France, and our findings may not apply to other settings or regions.

Author affiliations

¹Computational and Mathematical Biology Team, TIMC-IMAG UMR 5525, CNRS, Université Grenoble Alpes, Saint-Martin-d'Herès, Rhône-Alpes, France

²Clinical Epidemiology Unit, CHU Grenoble Alpes, Grenoble, Auvergne-Rhône-Alpes, France

³CIC 1406, INSERM, Université Grenoble Alpes, Saint-Martin-d'Herès, Rhône-Alpes, France

⁴Medical School, Université Grenoble Alpes, Saint-Martin-d'Herès, Rhône-Alpes, France

⁵Clinical Research Department, CHU Grenoble Alpes, Grenoble, Auvergne-Rhône-Alpes, France

Twitter Félix Pelen @fpelen

Acknowledgements The authors thank the research team in Grenoble Alpes University Hospital, without whose support this study could not be conducted. They also thank Nicolas Gonnet (CIC 1406, INSERM, Grenoble Alpes University, France) for the project monitoring and Jocelyne Deglaine (AGIL'OA company) for communication training. The authors thanks the participating physicians: Cécile Sabourdy, Sophie Blaise, Claire Cracowski, Fabienne Germe, Alban Deroux, Thomas Jouve, Juliette Coudurier, Maxime Lugosi, Emmeline Lagrange, Jean-Alexandre Long, Philippe Zaoui, Frédéric Garban, Mathilde Gallice, Olivier Heck, Caroline Makowski, Anne Brunet-Humbert, Olivier Ormezzano, Claude Bulaboïs, Marie Sicot, Sylvianne Hennebicq, Pascale Hoffmann, Céline Gallazzini and Françoise Sarrot-Reynaud.

Contributors AB conceptualised the study and is the guarantor. AB and PC developed the protocol and drafted the initial manuscript. AB and JL provided statistical expertise. AB, PC, AP, ZP, GC, SC and FP will be involved in the acquisition of data. All authors critically reviewed the protocol and approved submission of the final manuscript.

Funding This project is supported by Grenoble Alpes University Hospital (2019-A01621-56). The statistical analysis will be developed within the Grenoble Alpes Data Institute, which is supported by the French National Research Agency under the 'Investissements d'avenir' program (ANR-15-IDEX-02).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Alexandre Bellier <http://orcid.org/0000-0003-0907-0315>

José Labarère <http://orcid.org/0000-0001-7621-6586>

REFERENCES

- 1 Roter DL, Hall JA. Studies of doctor-patient interaction. *Annu Rev Public Health* 1989;10:163-80.
- 2 Schirmer JM, Mauksch L, Lang F, et al. Assessing communication competence: a review of current tools. *Fam Med* 2005;37:184-92.
- 3 Wensing M, Jung HP, Mainz J, et al. A systematic review of the literature on patient priorities for general practice care. Part 1: description of the research domain. *Soc Sci Med* 1998;47:1573-88.
- 4 Michie S, Miles J, Weinman J. Patient-centredness in chronic illness: what is it and does it matter? *Patient Educ Couns* 2003;51:197-206.
- 5 Candib LM. When the doctor takes off her clothes: reflections on being seen by patients in the exercise setting. *Families, Systems, & Health* 1999;17:349-63.
- 6 Junod Perron N, Sommer J, Louis-Simonet M, et al. Teaching communication skills: beyond wishful thinking. *Swiss Med Wkly* 2015;145:w14064.
- 7 Neumann M, Edelhäuser F, Tauschel D, et al. Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Acad Med* 2011;86:996-1009.
- 8 Zill JM, Christalle E, Müller E, et al. Measurement of physician-patient communication--a systematic review. *PLoS One* 2014;9:e112637.
- 9 Hulsman RL, Ros WJ, Winnubst JA, et al. Teaching clinically experienced physicians communication skills. A review of evaluation studies. *Med Educ* 1999;33:655-68.
- 10 Sanson-Fisher R, Hobden B, Waller A, et al. Methodological quality of teaching communication skills to undergraduate medical students: a mapping review. *BMC Med Educ* 2018;18:151.
- 11 Buckley S, Ambrose L, Anderson E, et al. Tools for structured team communication in PRE-REGISTRATION health professions education: a best evidence medical education (BEME) review: BEME guide No. 41. *Med Teach* 2016;38:966-80.
- 12 Stein T, Frankel RM, Krupat E. Enhancing clinician communication skills in a large healthcare organization: a longitudinal case study. *Patient Educ Couns* 2005;58:4-12.
- 13 Fossli Jensen B, Gulbrandsen P, Dahl FA, et al. Effectiveness of a short course in clinical communication skills for hospital doctors: results of a crossover randomized controlled trial (ISRCTN22153332). *Patient Educ Couns* 2011;84:163-9.
- 14 Zolnierok KBH, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care* 2009;47:826-34.
- 15 Gulbrandsen P, Jensen BF, Finset A, et al. Long-Term effect of communication training on the relationship between physicians' self-efficacy and performance. *Patient Educ Couns* 2013;91:180-5.
- 16 Irving G, Neves AL, Dambha-Miller H, et al. International variations in primary care physician consultation time: a systematic review of 67 countries. *BMJ Open* 2017;7:e017902.
- 17 Wilson AD, Childs S, Gonçalves-Bradley DC, et al. Interventions to increase or decrease the length of primary care physicians' consultation. *Cochrane Database of Systematic Reviews* : 2016;21.
- 18 Brown JB, Boles M, Moolooly JP, et al. Effect of clinician communication skills training on patient satisfaction. A randomized, controlled trial. *Ann Intern Med* 1999;131:822-9.
- 19 Price PB, Lewis EG, Loughmiller GC, et al. Attributes of a good practicing physician. *J Med Educ* 1971;46:229-37.
- 20 Hashim MJ. Patient-Centered communication: basic skills. *Am Fam Physician* 2017;95:29-34.
- 21 Bellier A, Chaffanjon P, Krupat E, et al. Cross-Cultural adaptation of the 4-Habits coding scheme into French to assess physician communication skills. *PLoS One* 2020;15:e0230672.
- 22 Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687.
- 23 Schulz R, Czaja SJ, McKay JR, et al. Intervention taxonomy (ITAX): describing essential features of interventions. *Am J Health Behav* 2010;34:811-21.
- 24 Krupat E, Frankel R, Stein T, et al. The four habits coding scheme: validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns* 2006;62:38-45.
- 25 Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol* 2015;68:435-41.
- 26 Webster G. *Final Report on the Patient Satisfaction Questionnaire Project*. Philadelphia, PA: American Board of Internal Medicine Committee on Evaluation of Clinical Competence, 1989.
- 27 Baillargeon P, Leduc A. *Inventaire de l'alliance thérapeutique (IAT)*. Trois-Rivières, QC: Université du Québec, Trois-Rivières, 2002.
- 28 Dion G, Tessier R. Validation de la traduction de l'Inventaire d'épuisement professionnel de Maslach et Jackson. [Validation of a French translation of the Maslach Burnout Inventory (MBI)]. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement* 1994;26:210-27.
- 29 Donner A. Approaches to sample size estimation in the design of clinical trials--a review. *Stat Med* 1984;3:199-214.
- 30 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377-99.
- 31 Chan A-W, Tetzlaff JM, Altman DG, et al. Spirit 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200-7.
- 32 Campbell MK, Piaggio G, Elbourne DR, et al. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661.

4.3. Partie 9) Résultats préliminaires de l'étude EPECREM

Dans ce travail préliminaire, nous avons cherché à explorer les premiers résultats de l'étude EPECREM. La pandémie de COVID-19 a retardé le calendrier initial de l'étude et ne permet pas de présenter l'intégralité des résultats à ce stade. Toutefois, les inclusions de patients se sont terminées à l'été 2021 et les évaluations des compétences relationnelles sont terminées depuis janvier 2022. Le plan d'analyse statistique doit encore être finalisé et l'analyse statistique sera réalisé en aveugle du groupe de randomisation par un statisticien indépendant. Sur la base des données recueillies, nous proposons en attendant une analyse descriptive des résultats de l'étude, pour l'ensemble des critères de jugement de l'étude. Nous avons ainsi choisi de ne pas présenter les résultats de l'analyse univariée ou multivariée à ce stade.

Nous avons envoyé une invitation par mail pour participer à l'étude aux 796 médecins éligibles exerçant au CHU Grenoble Alpes (Figure 1). 27 médecins ont alors manifesté leur intérêt pour participer à l'étude et 25 médecins ont été randomisés, certains n'ayant pas enregistré de consultation. Après randomisation, chaque médecin était invité à enregistrer 4 consultations en période pré-intervention et 4 consultations en période post-intervention.

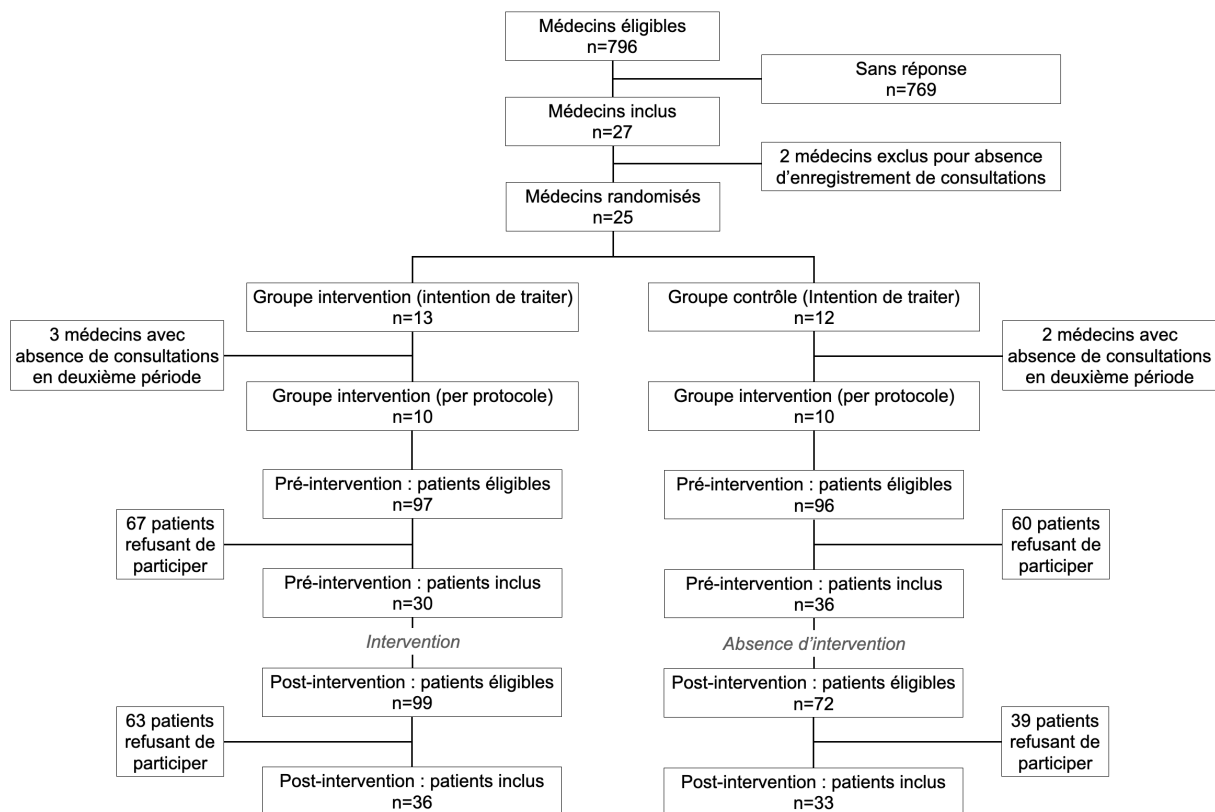


Figure 1. Diagramme de flux de l'étude EPECREM

66 patients, soit 33,6% des patients éligibles, ont accepté de participer à l'étude dans le groupe intervention contre 69 patients dans le groupe contrôle, soit 41,1% des patients éligibles.

Les médecins comprenaient 8 femmes (61,5%) dans le groupe intervention contre 6 (50,0%) dans le groupe contrôle. Le nombre de chirurgien et de titulaire sont à retrouver dans le Tableau 1.

Les patients inclus avaient en moyenne 49,8 ans (écart-type : 16,9) et 39 étaient des femmes (59,1%) dans le groupe intervention contre 53,8 ans (écart-type : 16,7 ans) dans le groupe contrôle avec 37 femmes (56,1%). L'ensemble des caractéristiques des patients sont présentées dans le Tableau 2.

		Groupe intervention (n=13)	Groupe contrôle (n=12)	Total (n=25)
Caractéristiques des médecins	Sexe féminin — n (%)	8 (61,5)	6 (50,0)	14 (56,0)
	Statut titulaire — n (%)	11 (84,6)	9 (75,0)	18 (72,0)
	Discipline chirurgicale — n (%)	1 (7,7)	2 (16,6)	3 (12,0)

Tableau 1. Caractéristiques des médecins

Caractéristiques des patients	Groupe intervention			Groupe contrôle			Total (n=135)
	Avant (n=30)	Après (n=36)	Total (n=66)	Avant (n=36)	Après (n=33)	Total (n=69)	
Age moyen (écart-type)	47,1 (15,5)	52,4 (18,4)	49,8 (16,9)	52,3 (17,9)	55,3 (14,9)	53,8 (16,7)	51,9 (16,9)
Sexe féminin — n (%)	22 (73,3)	17 (47,2)	39 (59,1)	20 (58,8)	17 (51,5)	37 (56,1)	76 (56,3)
Marié — n (%) [§]	15 (55,6)	25 (69,4)	40 (60,6)	17 (58,6)	20 (64,5)	37 (56,1)	77 (57,0)
Vit hors agglomération — n (%) [§]	16 (59,3)	10 (27,8)	26 (39,4)	16 (55,2)	11 (35,5)	27 (40,1)	53 (39,3)
Accompagnement par un tiers — n (%) [§]	9 (33,3)	11 (30,6)	20 (30,3)	9 (31,0)	12 (38,7)	21 (30,4)	41 (30,4)

Tableau 2. Caractéristiques des patients

[§] Missing data : n=10

Dans cet essai contrôlé randomisé, nous avons effectué des mesures des compétences relationnelles via l'échelle 4HCS durant la période pré-intervention et post-intervention pour chacun des groupes. Ainsi, la différence des différences pour le critère de jugement principal (score global moyen 4HCS) était de 6,8 points (Figure 2). Une progression du score de

compétences relationnelles était donc mesurée après l'intervention pour le groupe intervention, mais cette progression était d'autant plus importante qu'une baisse était observée sur la même période dans le groupe contrôle.

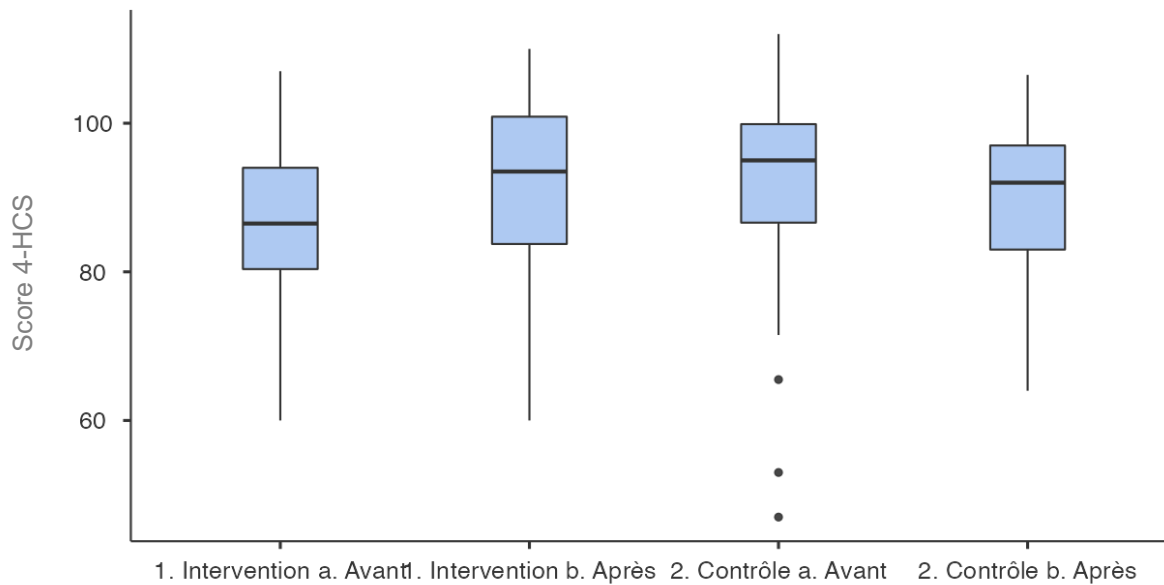


Figure 2. Graphique de l'évolution du score 4HCS en fonction du groupe et de la période

Selon les dimensions de l'échelle 4HCS, la différence des différences était de 2,4 points pour Habit 1 (S'investir dès le début), de 0,9 point pour Habit 2 (Obtenir le point de vue du patient), de 0,8 point pour Habit 3 (Faire preuve d'empathie) et de 2,7 points pour Habit 4 (S'investir jusqu'à la fin) (Figure 3). Une progression du score de compétences relationnelles était donc mesurée après l'intervention pour le groupe intervention quel que soit la dimension considérée, alors qu'une baisse était observée sur la même période dans le groupe contrôle.

Selon les dimensions de l'échelle 4HCS, la différence des différences était de 5,3 points pour l'alliance thérapeutique (échelle Inventory of the Therapeutic Alliance), de -2,6 points pour la satisfaction du patient (échelle American Board of Internal Medicine Patient Satisfaction Rating Scale), et de 5,8 points pour l'accomplissement personnel (échelle Maslach Burnout Inventory — MBI) (Figure 4). Seule la satisfaction connaissait une baisse dans le groupe intervention entre la période avant et après intervention (Tableau 2).

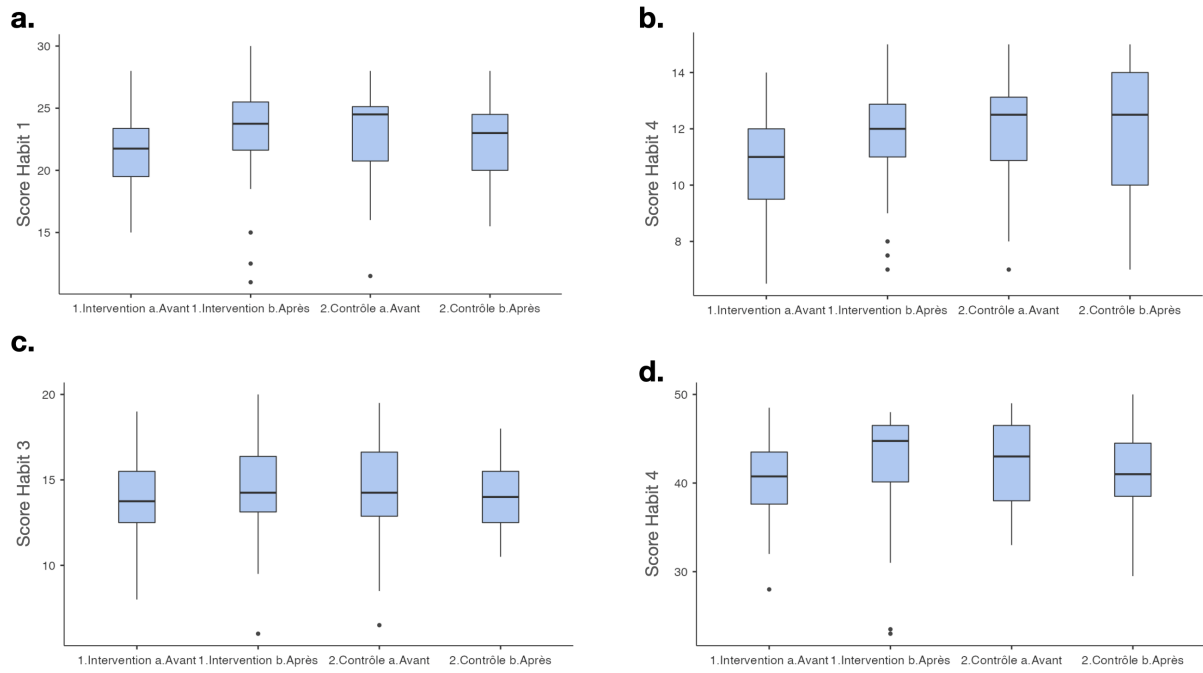


Figure 3. Graphique de l'évolution des dimensions du 4HCS en fonction du groupe et de la période.

a. Dimension Habit 1, **b.** Dimension Habit 2, **c.** Dimension Habit 3, **d.** Dimension Habit 4

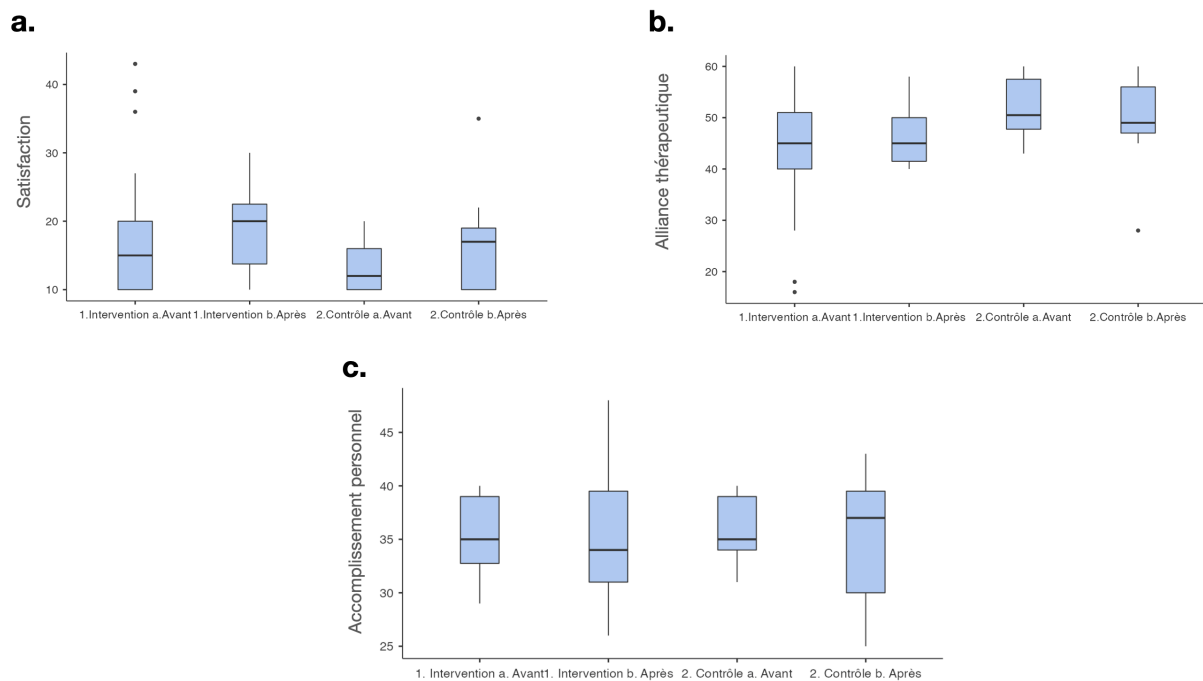


Figure 4. Graphique de l'évolution des critères secondaires en fonction du groupe et de la période.

a. Score d'alliance thérapeutique, **b.** Score de satisfaction, **c.** Score d'accomplissement personnel

SYNTHESE ET PERSPECTIVES

5. SYNTHÈSE ET PERSPECTIVES

5.1. Évaluation des compétences relationnelles médecin-patient

Bien qu'une évaluation standardisée des compétences relationnelles soit souhaitable, nous pouvons regretter le manque d'instruments disponibles. En particulier, notre revue systématique de la littérature a mis en évidence le manque d'instruments avec une validation satisfaisante des propriétés psychométriques. Il semble que la rigueur de la psychométrie soit limitée dans le champ des études en lien avec l'éducation médicale ou l'évaluation de compétences médicales non-techniques, à l'instar des autres domaines et malgré une progression de la qualité de la documentation ces dernières années. Les analyses factorielles ou le processus de génération des items sont rarement décrits. Ainsi, il est difficile d'appréhender la signification précise de la mesure permise par les échelles d'évaluation des compétences relationnelles.

Par ailleurs, il a été difficile d'identifier des instruments de mesure de ces compétences, telles que nous les avons définis. Seules 9 échelles d'évaluation ont été retrouvées. C'est un nombre nettement inférieur à ce que des revues de la littérature ont retrouvé pour des échelles d'évaluation du seul concept de communication, avec une soixantaine d'échelles identifiées (Boon and Stewart 1998; Zill et al. 2014).

Rares sont aussi les échelles d'évaluation ayant été utilisées dans des contextes différents, à la fois en formation initiale et en formation continue, pour de l'évaluation et pour de la formation, avec des patients réels et des patients standardisés (Boon and Stewart 1998).

L'échelle 4-Habits Coding Scheme (4-HCS), identifiée lors de notre revue systématique, répond à l'ensemble de ces critères. C'est une échelle largement utilisée à travers le monde depuis plus de 15 ans (Stein et al. 2005; Krupat et al. 2006; Gulbrandsen et al. 2013) d'abord dans un contexte hospitalier, puis dans un contexte universitaire de simulation (Bosméan et al. 2022). C'est une échelle qui permet l'évaluation des compétences de communication verbale et non verbale, ainsi que de l'empathie, de la relation centrée sur le patient ou de la structuration de l'entretien, recouvrant ainsi l'ensemble du champ des compétences relationnelles médecin-patient (Krupat et al. 2006). Enfin, c'est une échelle dont nous avons validé les propriétés psychométriques selon les standards internationaux (Nunnally and Bernstein 1994), après adaptation trans-culturelle réalisée par une société spécialisée de traduction (Epstein et al. 2015).

Cette adaptation trans-culturelle a permis de disposer d'une version en langue française, après l'anglais (Krupat et al. 2006), l'allemand (Scholl et al. 2014) ou le norvégien (Fossli Jensen et al. 2010). Notre étude de validation a permis de confirmer les propriétés satisfaisantes de l'échelle concernant la cohérence interne, avec un excellent coefficient Alpha de Cronbach à 0,94 par exemple, et la fiabilité. Concernant cette dernière, nous avons toutefois mis en évidence que pour une mesure satisfaisante, il valait mieux considérer une double évaluation indépendante. Seule cette double évaluation permettait de dépasser le seuil usuel de 0,70 pour le coefficient de corrélation intra-classe (Terwee et al. 2007), avec un coefficient mesuré alors à 0,72. Pour les études suivantes utilisant l'échelle 4-HCS, nous avons donc effectué une évaluation par deux médecins experts indépendants, en aveugle.

Notre étude de validation de l'échelle 4-HCS a également démontré pour la première fois la structure dimensionnelle de l'instrument (Bellier et al. 2020), remettant en question la structure en quatre dimensions de l'étude originale (Krupat et al. 2006). En effet, l'instrument de mesure semble plutôt bi-dimensionnel avec une dimension empirique recoupant les items de communication et de structuration de l'entretien et une autre dimension recoupant les items de relation centrée sur le patient (Bellier et al. 2020). Nous pourrions donc discuter la pertinence des sous-scores par dimensions (Habits 1 à 4) qui sont largement rapportés dans les études utilisant l'échelle 4-HCS.

Enfin, la validation de l'échelle 4-HCS a été réalisée dans des contextes différents de ceux des précédentes études hospitalières (Krupat et al. 2006; Fossli Jensen et al. 2010), puisque nous avons d'abord validé l'échelle pour la formation initiale avec des simulations utilisant des patients standardisés (Bellier et al. 2020), puis pour d'autres professionnels médicaux qui n'avaient pas encore été étudiés via cette échelle : les sages-femmes (Blanc et al. 2020) et les médecins généralistes libéraux. Nous avons alors mis en évidence une bonne transposabilité des propriétés psychométriques de la version française de l'échelle 4-HCS en fonction du contexte d'utilisation : patients standardisés versus patients réels, secteur hospitalier versus ambulatoire ou encore spécialistes versus généralistes (Tableau 4). Le coefficient alpha de Cronbach global reste supérieur à 0,90 et la fiabilité supérieure à 0,60. La validation de la version française du 4-HCS a été étudiée sur 481 consultations différentes, qui s'ajoute aux 666 consultations précédentes dans d'autres pays. L'échelle 4-HCS a par ailleurs une bonne transposabilité par rapport aux résultats obtenus dans la version originale ou dans les autres adaptations trans-culturelles, renforçant ainsi la validité de l'échelle pour l'évaluation des compétences relationnelles médecin-patient (Tableau 4).

Études (année)	Patients				Médecins				Consistance interne		Fiabilité		
	Pays	Type	Nombre	Hommes n (%)	Age moyen	Type	Nombre	Hommes n (%)	Age moyen	Nombre de consultations	Alpha de Cronbach	Corrélation inter-items	Fiabilité inter- juge*
Étude EPECREM [§] (2022)	France	Réels	135	57 (32,8)	48,8	Spécialistes hospitaliers	25	14 (56)	135	0,95	0,67	0,61	0,99
Étude médecine générale [†] (2021)	France	Réels	88			Généralistes libéraux	10	5 (50)	88	0,95	0,69	0,82	0,99
Blanc (2020)	France	Réels	58	0 (0)		Sages- femmes	12	0 (0)	58	0,97	0,60	0,72	0,96
Bellier (2020)	France	Standardisés	20			Étudiants en médecine	200		200	0,94	0,42	0,60	0,72
Scholl (2014)	Allemagne	Réels	67	25 (37)	55,7	Généralistes, spécialistes salariés	22	12 (55)	69	0,44	0,40	0,77	0,72
Fossli Jensen (2010)	Norvège	Réels	497	239 (48)	46,0	Spécialistes hospitaliers	71	55 (62)	497			0,74	
Krupat (2006)	États-Unis	Réels	100	25 (50)	62,0	Spécialistes hospitaliers et internes	50	25 (50)	100	0,66		0,72	

Tableau 4. Caractéristiques de l'échantillon et de l'outil de mesure 4-HCS parmi les études de validation de cette échelle

*Coefficient de corrélation intra-classe basé sur les évaluations individuelles sauf pour Krupat et al. (coefficient de corrélation de Pearson)

[§]Voir Partie 9) Résultats préliminaires de l'étude EPECREM

[†] Voir Partie 3) Compétences relationnelles en médecine générale ambulatoire

A partir d'un outil d'évaluation valide et fiable des compétences relationnelles médecin-patient, la question du contexte d'utilisation de cette échelle se pose.

L'utilisation de patients standardisés pour l'évaluation des compétences des étudiants en médecine semble fiable dans la littérature anglo-saxonne (van Zanten et al. 2007), tout comme la mesure effectuée directement en stage par l'observation au lit du patient (Jouriles et al. 2002). Toutefois, il convient de différencier la mesure des compétences relationnelles, de l'impact direct sur le ressenti du patient, en particulier sur sa satisfaction. Ainsi, lors d'Examens Cliniques Objectifs Structurés (ECOS) faisant appel à ces patients, la corrélation entre les résultats aux ECOS et la satisfaction du patient standardisé n'a pas été établie (Chessman et al. 2003). Dans notre expérience, nous retrouvons des résultats similaires. En effet, lorsque nous avons demandé aux patients standardisés d'évaluer leur satisfaction à l'issue de la consultation, les résultats étaient différents de ceux obtenus avec l'échelle 4-HCS. Nos étudiants qui avaient bénéficié d'une formation avaient des compétences relationnelles significativement augmentées, alors que l'augmentation de la satisfaction des patients était faible et non significative (Bosméan et al. 2022). Dans les résultats préliminaires de notre essai contrôlé randomisé (étude EPECREM), nous avons également retrouvé la même tendance. Entre le groupe intervention et le groupe contrôle, la satisfaction du patient est le seul critère de jugement qui diminue malgré la formation des médecins hospitaliers.

La recherche systématique de la satisfaction du patient pourrait donc être une mesure intéressante et complémentaire pour approcher de manière plus fine les compétences relationnelles et leur impact. Des échelles d'évaluation existent, de la plus simple (Donnelly et al. 2000) à la plus complexe, mais validée en langue française (Morel et al. 2018), pour permettre le retour du patient sur la consultation.

Par ailleurs, il peut exister des biais dans la mesure des compétences relationnelles du fait de l'évaluateur lui-même. Dans notre étude de validation, nous avons ainsi mis en évidence que les évaluateurs les plus expérimentés étaient ceux qui attribuaient les scores les plus hauts. A contrario, les étudiants en médecine qui évaluaient leurs pairs étaient ceux qui attribuaient les scores les plus bas, du fait d'une application plus stricte de l'échelle d'évaluation 4-HCS (Bellier et al. 2020). Toutefois, la reproductibilité des mesures restait satisfaisante quelle que soit la paire d'évaluateurs.

L'évaluation des compétences relationnelles pourrait également être influencée par le sexe de l'évaluateur (Cuddy et al. 2011). Sur la base de résultats post-hoc tirés nos études, nous ne retrouvons pas de différence significative ($p=0,766$) entre le score moyen des femmes et des hommes, respectivement 85,1 (écart-type : 19,6) sur 219 consultations et 84,6 (écart-type :

16,5) sur 532 consultations. Ce lien semble donc encore incertain et mériterait d'être étudié de manière plus approfondie.

Dans ce contexte, une auto-évaluation à visée formative pourrait être envisagée en complément de l'hétéro-évaluation. L'échelle 4-HCS n'a été que rarement utilisée en auto-évaluation. Toutefois, dans le cadre de l'essai EPECREM (Bellier et al. 2022), nous avons encouragé les médecins du groupe intervention à utiliser cet outil pour mesurer leurs compétences relationnelles durant les consultations. Par ailleurs, sur la base de consultations vidéo-enregistrées, l'auto-évaluation des compétences relationnelles a déjà été démontrée comme faisable (Zick et al. 2007).

5.2. Déterminants des compétences relationnelles en consultation médicale

A partir de l'échelle 4-HCS, nous avons pu étudier les déterminants multiples d'une consultation médicale réussie sur le plan de la relation médecin-patient. Tout d'abord, nous avons noté des différences importantes dans les scores obtenus en fonction du professionnel étudié (Tableau 5). Ainsi, un premier déterminant majeur semble être la profession et le secteur d'activité. Les médecins généralistes étudiés avaient des compétences relationnelles (score global à l'échelle 4-HCS de 84,1) supérieures à celles mesurées dans la littérature, principalement constituée de médecins hospitaliers. Cependant, les médecins hospitaliers français que nous avons inclus dans le cadre de l'étude EPECREM (Bellier et al. 2022) avaient des scores supérieurs (score global à l'échelle 4-HCS de 89,9) et les sages-femmes ont eu les meilleurs scores (score global à l'échelle 4-HCS de 98,7) (Blanc et al. 2020). Les étudiants en médecine avaient logiquement les scores les moins élevés (score global à l'échelle 4-HCS de 76,4) (Bellier et al. 2020).

Si nous nous intéressons aux différentes dimensions (Habits 1 à 4) de l'échelle 4-HCS, nous retrouvons les mêmes tendances (Tableau 5). Cependant, les étudiants en médecine peuvent avoir des scores supérieurs sur les dimensions de relation centrée sur le patient (c'est-à-dire, Obtenir le point de vue du patient) et d'empathie (c'est-à-dire, Faire preuve d'empathie) par rapport aux médecins hospitaliers et aux médecins généralistes (Bosméan et al. 2022). Ces deux dimensions nécessitent sans doute moins d'expérience clinique que les deux autres étant avant tout des compétences humaines et sociales. La prise en charge du patient et l'explication du plan thérapeutique en alliance avec le patient sont souvent réservés à un apprentissage plus tardif au cours du cursus médical, généralement dans le troisième cycle des études. Cela

explique que les médecins séniors aient de meilleurs résultats sur la dernière dimension (c'est-à-dire, S'investir jusqu'à la fin).

Ce constat doit nous encourager à débiter de manière précoce la sensibilisation au relationnel médecin-patient (Rees et al. 2002) et l'évaluation des étudiants en santé durant leur cursus (Zick et al. 2007).

Études (années)	Score Habit 1 <i>moyenne</i> (<i>écart-type</i>)	Score Habit 2 <i>moyenne</i> (<i>écart-type</i>)	Score Habit 3 <i>moyenne</i> (<i>écart-type</i>)	Score Habit 4 <i>moyenne</i> (<i>écart-type</i>)	Score global <i>moyenne</i> (<i>écart-type</i>)
Étude					
EPECREM [§] (2022)	22,5 (3,7)	11,6 (2,1)	14,1 (2,7)	41,7 (5,4)	89,9 (12,3)
Étude médecine générale [†] (2021)	21,4 (5,1)	10,4 (2,6)	14,1 (3,3)	38,2 (7,6)	84,1 (16,6)
Blanc (2020)	26,9 (3,7)	11,9 (2,9)	16,9 (3,4)	43 (6,6)	98,7 (15,7)
Bellier (2020)	18,8 (2,8)	10,0 (2,1)	13,8 (3,0)	33,9 (6,1)	76,4 (12,3)
Scholl (2014)	12,1 (5,4)	4,6 (2,7)	3,6 (1,9)*	26,5 (11,1)	46,8 (21,1)*
Fossli Jensen (2010)					60,2 (15,0)
Krupat (2006)	17,6 (6,6)	7,6 (3,7)	11,3 (4,9)	31,5 (10,1)	68,0 (25,3)

Tableau 5. Scores de l'échelle 4-HCS dans les différentes études de validation

[§] Voir Partie 9) Résultats préliminaires de l'étude EPECREM

[†] Voir Partie 3) Compétences relationnelles en médecine générale ambulatoire

* L'item 13 n'a pas pu être évalué dans cette étude

Un autre déterminant des compétences relationnelles que nous avons mis en évidence est le sexe du médecin. Quel que soit l'échantillon étudié, les femmes ont systématiquement eu des scores significativement plus élevés que leurs homologues masculins à l'échelle 4-HCS. La différence est d'autant plus importante chez les médecins généralistes inclus et elle peut sans doute en partie expliquer les meilleurs résultats des sages-femmes, les professionnels de la

maïeutique que nous avons inclus étant tous féminins. Ces résultats sont en cohérence avec la littérature comme nous l'avons évoqué précédemment (Roter et al. 2002).

En complément de ce facteur déjà identifié, nous avons mis en évidence l'absence de lien statistique entre les résultats aux examens théoriques d'un étudiant en médecine et ses compétences relationnelles. Cette absence de lien renforce l'intérêt d'une évaluation complète des étudiants en médecine au cours de leur cursus, axée à la fois sur les connaissances et sur les compétences, les meilleurs étudiants aux examens n'étant pas toujours les meilleurs aux ECOS (Adam et al. 2012).

Notre étude a également montré l'intérêt d'une plus grande ouverture et d'une plus grande diversité des parcours au cours du deuxième cycle des études médicales, puisque les étudiants ayant effectué des stages à l'étranger ou des stages en laboratoire de recherche ont des compétences relationnelles significativement augmentées. Ce résultat va dans le sens de la réforme du deuxième cycle des études de médecine en France (Arrêté du 21 décembre 2021 relatif à l'organisation des épreuves nationales donnant accès au troisième cycle des études de médecine), qui prévoit la prise en compte du parcours de l'étudiant, au-delà de son investissement universitaire classique. Cependant, ce résultat peut interroger car la réalisation de tels stages en dehors des Centres Hospitaliers Universitaires engendre une diminution du nombre d'heures en immersion clinique, principalement pour les stages en laboratoire de recherche. Cela doit nous interroger sur l'apprentissage des compétences lors des stages hospitaliers, qui est peut-être plus centré sur les compétences techniques que les compétences non-techniques. Il faut certainement sensibiliser les enseignants et les médecins accompagnant les externes en médecine dans les services pour qu'ils puissent effectuer plus régulièrement une évaluation des compétences relationnelles par observation directe de l'étudiant auprès du patient en chambre (Jouriles et al. 2002).

Enfin, le dernier déterminant identifié des compétences relationnelles médecin-patient est la durée de consultation. Répondant à une problématique soulevée par les cliniciens, nous avons montré que la durée de consultation augmentait avec les compétences relationnelles. Cela explique sans doute en partie les résultats plus faibles que nous avons obtenus pour les médecins généralistes avec une durée moyenne de consultation de 14 minutes, par rapport aux médecins hospitaliers qui avaient dans l'étude EPECREM une durée de consultation moyenne de 30 minutes. Toutefois, la relation entre la durée de consultation et le score 4-HCS n'était pas linéaire. Avec un seuil autour de 13 minutes, à partir duquel l'augmentation du score 4-HCS de compétences relationnelles devient faible pour chaque minute de consultation supplémentaire,

nous avons montré qu'il était possible d'avoir un bon niveau de compétences relationnelles avec une consultation d'une durée contenue. Ce résultat est consolidé par l'observation de la même tendance chez les étudiants en médecine qui avaient pourtant une durée de consultation limitée par l'exercice de simulation. Bien entendu, le motif de consultation modifie de manière importante la durée de consultation. Bien que les maladies graves comme les cancers semblent entraîner de la part du médecin une mise en jeu de meilleures compétences, les écarts observés en termes de compétences relationnelles ne permettent pas à ce stade de dégager des motifs de consultation qui entraîneraient un relationnel médecin-patient significativement augmenté.

5.3. Développement des compétences relationnelles par des programmes pédagogiques dédiés

Les déterminants des compétences relationnelles identifiés au cours de nos différentes études montrent finalement qu'une part importante est modifiable par le contenu de la formation initiale ou le développement des compétences en formation continue. L'enjeu des compétences relationnelles médecin-patient à ce stade se situe donc dans la formation et la mise en place de programmes pédagogiques dédiés à leur développement.

Pour optimiser ces compétences non-techniques, il est possible de s'inspirer des nombreuses interventions décrites dans la littérature (Gilligan et al. 2021). Avec 76 interventions dédiées à la communication pour les étudiants en médecine, dont une majorité d'essais contrôlés randomisés et une majorité de résultats positifs sur divers critères de jugement, le choix est vaste. Cependant, les études s'étant intéressées au développement des compétences relationnelles au sens large sont plus rares, tout comme les interventions déclinées en formation continue, en complément de la formation initiale des médecins (Gilligan et al. 2021).

Quelques modèles pédagogiques ont malgré tout émergé, en partant généralement d'un outil de mesure des compétences communicationnelles pour proposer des méthodes éducatives permettant de développer spécifiquement chacun des items. Le modèle de Calgary-Cambridge fait partie de ces modèles dominants (Kurtz and Silverman 1996), tout comme la boîte à outils du Conseil d'accréditation pour l'enseignement supérieur en médecine des États-Unis (Rider and Keefer 2006).

Nous avons choisi dans un premier temps d'évaluer les interventions pédagogiques qui s'étaient progressivement mises en place au sein de la Faculté de Médecine de l'Université Grenoble

Alpes ou au sein de structures associatives locales, destinées à un public d'étudiants en médecine (Bosméan et al. 2022). Ces interventions, mises en place de manière empirique la plupart du temps, reposaient généralement sur des approches pédagogiques multiples et recoupaient en grande partie les modèles cités précédemment (Kurtz and Silverman 1996; Rider and Keefer 2006). Elles se sont montrées par ailleurs efficaces sur les compétences relationnelles pour la majorité d'entre elles, avec une efficacité accrue de manière linéaire avec l'augmentation du nombre de formations suivies. Cette étude démontre que le potentiel d'une formation aux compétences relationnelles ne réside pas tant dans le contenu ou le format choisi, mais surtout dans sa répétition. Par ailleurs, l'effet semblait persister dans le temps, en particulier lorsque les étudiants ont été confrontés dès le début de leur cursus à ce type de formation (Zick et al. 2007; Bosméan et al. 2022).

En formation initiale, plusieurs points de vigilance sont à noter. Il convient de prêter une attention particulière aux cas cliniques des patients standardisés qui peuvent influencer le niveau de l'évaluation des compétences relationnelles. Il est nécessaire de standardiser et de catégoriser les scénarii en fonction de l'objectif d'apprentissage (Bank et al. 2021).

Par ailleurs, il convient de prendre en compte les obstacles aux programmes de formation. Elles sont multiples et concernent tout d'abord des aspects organisationnels, comme le manque de temps dédié, le manque de valorisation de ces activités d'enseignement, la mise en concurrence avec d'autres activités techniques ou scientifiques ou la complexité ainsi que le coût de mise en œuvre lors de simulations (Yudkowsky et al. 2006; Junod Perron et al. 2015). Elles concernent également l'attitude parfois négative des médecins sur l'enseignement de ces compétences non-techniques avec de fait un mauvais modèle pour l'étudiant (Junod Perron et al. 2015). Enfin, les freins peuvent concerner les étudiants eux-mêmes avec des attitudes parfois négatives sur le travail des compétences relationnelles. Ces dernières peuvent être considérées comme des compétences qui ne sont pas du ressort de l'université, comme des compétences de bon sens ou comme des compétences moins sérieuses, plus subjectives que les compétences techniques (Rees et al. 2002). Ainsi, un enjeu de la réussite des formations aux compétences relationnelles est de travailler sur la représentation de ces compétences et de mieux valoriser la relation humaine médecin-patient au cours du cursus, en particulier lors du deuxième cycle. Une des solutions pourrait être de favoriser l'implication des étudiants et de mettre en place des évaluations systématiques sur les formations mises en place (Rees et al. 2004). Lorsque l'avis des étudiants est recueilli, il ressort que les formats interactifs, en particulier de jeux de rôle, sont plébiscités pour la formation au relationnel (Rees et al. 2004).

D'autres formats alternatifs pourraient également avoir un intérêt pour susciter l'adhésion des étudiants, comme les outils modernes de réalité virtuelle. Ces derniers ont montré leur efficacité en particulier pour les étudiants en difficulté, grâce à l'environnement immersif et la possibilité de répéter l'exercice à l'envie (Kleinlogel et al. 2021). Au contraire, un rapprochement avec les sciences humaines pourrait aussi être bénéfique pour l'amélioration des compétences relationnelles, en particulier le travail de l'étudiant sur lui-même. Les techniques de focalisation sur la relation intérieure ou la méditation peuvent ainsi améliorer l'écoute empathique des étudiants (Alexander et al. 2015).

Après avoir étudié les interventions mises en place en formation initiale, nous avons décidé de construire une intervention autour de l'échelle 4-HCS (Stein et al. 2005; Krupat et al. 2006) pour le développement des compétences relationnelles de médecins séniors, en formation continue. Nous avons choisi de réaliser une intervention courte, en cohérence avec la faible disponibilité des médecins hospitaliers, et d'ajouter une introduction à des techniques de communication (Process Communication) avec une partie en jeux de rôle. Par ailleurs, un retour individuel sur les compétences relationnelles des médecins participants était réalisé par des experts de l'échelle 4-HCS. Pour cela, les vidéos des consultations en question étaient transmises avec les points forts et les points d'amélioration des compétences relationnelles (Bellier et al. 2022).

L'évaluation de cette intervention a été réalisée par une étude de haut niveau de preuve (c'est-à-dire un essai contrôlé randomisé). Les résultats préliminaires de cette étude semblaient montrer une amélioration des compétences relationnelles dans le groupe intervention de 6,8 points à l'échelle 4-HCS, alors que les scores des médecins du groupe contrôle diminuaient entre les deux périodes d'étude. L'augmentation des compétences relationnelles avec l'intervention s'observait sur toutes les dimensions de l'échelle et elle s'observait également sur l'alliance thérapeutique ou l'accomplissement personnel du médecin. Seule la satisfaction du patient ne montrait pas d'augmentation.

Les enseignements de cette étude sont, de fait, multiples pour l'amélioration du relationnel en formation continue. Tout d'abord, il convient de lutter contre l'érosion des compétences relationnelles au cours du temps et de la carrière d'un praticien. Cela doit passer par la réactivation des compétences, la re-sensibilisation aux problématiques de communication et la répétition des formations (Junod Perron et al. 2015). De plus, pour améliorer le dernier critère manquant à l'étude EPECREM, il convient de réfléchir à l'intégration du patient dans

l'évaluation ou la formation du médecin. Bien que les preuves restent limitées à ce jour concernant l'effet du retour du patient sur les compétences relationnelles du médecin (Cheraghi-Sohi and Bower 2008), il semble que l'implication des patients puisse améliorer l'empathie et certaines compétences spécifiques (Perry et al. 2013). Le cadre et le format de ces interventions restent toutefois à définir de manière plus précise.

D'après notre expérience de l'analyse de plusieurs programmes pédagogiques, nous suggérons quelques conseils opérationnels pour la mise en œuvre d'une formation (Tableau 6), avec une systématisation de la description de l'intervention sur le modèle des recommandations TIDieR (Hoffmann et al. 2014). La formation idéale pourrait être une intervention multifacette associant du contenu théorique, applicatif et pratique, suivi de retour individuel et personnalisé sur les compétences relationnelles. Ces interventions devraient être dispensées à plusieurs avec un médecin et un expert de la communication, en petit groupe et de manière répétée (Tableau 6).

5.4. Perspectives

L'évaluation, la compréhension et le développement des compétences relationnelles médecin-patient sont des enjeux importants pour les médecins. Nous avons suggéré qu'il était possible de les évaluer de manière fiable, qu'il était possible d'identifier des médecins ou des étudiants en médecine pouvant bénéficier préférentiellement de formation et que des formations dédiées pouvaient être efficaces. Ainsi, les universités françaises portent cette responsabilité de formation tout au long de la carrière, avec un équilibre difficile à trouver entre la formation scientifique théorique, l'apprentissage des compétences techniques et le développement de compétences non-techniques, l'ensemble étant nécessaire pour qu'un patient ait confiance dans un professionnel (Murakami et al. 2010). Cet équilibre passe par une médecine qui se recentre, à mi-chemin entre la biologie de Claude Bernard, qui a profondément influencé le développement médical au vingtième siècle (Olmsted 1935), et les sciences humaines avec le développement des humanités médicales (Dennhardt et al. 2016; Ilcewicz et al. 2018). Au-delà de la responsabilité des universités, il ne faut pas oublier la responsabilité individuelle des médecins et celle des établissements de santé. En effet, les compétences relationnelles ont un fort impact sur le patient et sur les soins. Elles constituent dès lors un indicateur de performance pour un établissement de santé ou pour un médecin libéral, pouvant aussi bien contribuer au prestige du praticien concerné qu'au rejet des patients de ce dernier (Chichirez and Purcărea 2018). Par extension, les compétences relationnelles développées avec le patient sont également

Critères	Propositions
Quoi ?	Théorie Importance des compétences relationnelles pour le soin, définition et limites des compétences relationnelles, grandes étapes de la consultation pour établir une relation de confiance via l'échelle 4-HCS et présentation des modèles de communication
	Application Visionnage d'extraits de consultations vidéo-enregistrées puis débriefing des points forts et des points d'amélioration et identification des profils de communication. Possibilité d'évaluation par l'échelle 4-HCS d'une consultation
	Mise en pratique Jeux de rôle en trinôme avec un médecin, un patient et un observateur puis changement des rôles avec scénarii de difficulté croissante (augmentation de l'état de stress du patient). Identification des points forts, des canaux de communication et discussion sur les points d'amélioration. Puis possibilité de jeux de rôle sur des situations courtes en groupe complet avec identification des points forts par l'échelle 4-HCS
	Feedback Analyse détaillée d'une ou de plusieurs consultations vidéo-enregistrées (réelles ou simulées) via l'échelle 4-HCS avec commentaire précisant les points forts et les pistes d'amélioration. Transmission de la vidéo à l'apprenant
Qui ?	Deux à trois intervenants dont un personnel médical expert et un personnel extérieur (psychologue, coach en communication, enseignant en sciences humaines) qui peut être associé ou remplacé par un patient expert
Comment ?	Formation en groupe d'une douzaine d'apprenants Salle facilement accessible, équipée de matériel audiovisuel et adaptée à l'interactivité (pas ou peu de table, disposition en cercle des apprenants)
Où ?	
Quand et combien de fois ?	Durée de 3 à 4 heures maximum pour chaque session, mais graduer la formation avec une répétition dans le temps avec un délai d'au minimum un mois, voire tous les ans pour les étudiants
Contenu adapté ?	Possibilité d'adaptation du contenu de la formation au niveau préalable des apprenants par évaluation des compétences relationnelles en amont ou auto-questionnaire pour déterminer le profil de communication

Tableau 6. Propositions pour construire une intervention pédagogique efficace

utiles pour les relations inter-professionnelles, qui ont un impact fort sur le soin (Ajam et al. 2017). Enfin, le travail des compétences relationnelles s'inscrit dans une réflexion internationale sur la certification des médecins. Faisant partie des six compétences ciblées par le Conseil d'accréditation américain pour l'enseignement supérieur en médecine, les compétences relationnelles et de communication pourraient faire l'objet d'une évaluation et de formations pour permettre une autorisation d'exercice, que ce soit en fin de formation initiale ou à intervalles réguliers, durant la formation continue (Heffron et al. 2007).

Une des limites au développement des compétences relationnelles est la forte consommation de ressources humaines et le caractère chronophage de l'évaluation. Aidée par les techniques de machine learning, une automatisation de l'évaluation pourrait être étudiée à l'avenir. Dans le cadre d'un partenariat avec un laboratoire de recherche public-privé grenoblois, spécialisé en intelligence artificielle, nous réfléchissons désormais à une cotation automatique des items de l'échelle 4-HCS. Sur la base des retranscriptions audio par reconnaissance vocale ou de l'analyse du signal de la voix, des algorithmes pourraient utiliser notre large base de données de près d'un millier de consultations vidéo-enregistrées pour l'apprentissage puis la validation de la cotation. Un tel outil pourrait permettre de réduire considérablement le temps enseignant en formation initiale, de multiplier les exercices d'entraînement avec auto-évaluation pour l'étudiant et de réduire la variabilité inter-évaluateur. Par ailleurs, cet outil pourrait être implémenté dans les logiciels médicaux à l'hôpital pour une analyse en continue des performances relationnelles, afin de permettre au médecin d'adopter des corrections au cours des consultations. La technique utilisée pourrait également être déclinée dans d'autres situations d'évaluation que les compétences relationnelles, dans de nombreux domaines liés aux soins.

5.5. Conclusion

Pour conclure, ce travail de thèse nous a permis d'explorer les compétences relationnelles médecin-patient, compétences majeures pour l'exercice médical et la qualité des soins. Les propriétés psychométriques des échelles d'évaluation de ces compétences restaient peu étudiées, tout comme les déterminants des compétences relationnelles. Par ailleurs, les études évaluant des programmes de formation aux compétences relationnelles étaient généralement de faible niveau de preuve. Nous avons donc identifié une échelle d'évaluation standardisée des compétences relationnelles médecin-patient grâce à une revue systématique de la littérature, puis nous avons effectué son adaptation trans-culturelle en langue française et la validation des

propriétés psychométriques de cet outil de mesure. Une évaluation fiable des compétences relationnelles a permis de comprendre les déterminants des compétences relationnelles, chez les étudiants en médecine ou les professionnels médicaux en exercice, afin de mieux appréhender l'efficacité des programmes pédagogiques. Nous avons ainsi évalué les formations dispensées par l'université aux étudiants en médecine, puis nous avons organisé l'évaluation par une étude à haut niveau de preuve d'un programme pédagogique multifacette de développement des compétences relationnelles des médecins hospitaliers.

6. Références bibliographiques

Adam J, Bore M, McKendree J, Munro D, Powis D. 2012. Can personal qualities of medical students predict in-course examination success and professional behaviour? An exploratory prospective cohort study. *BMC Med Educ.* 12:69. <https://doi.org/10.1186/1472-6920-12-69>

Aiarzaguena JM, Grandes G, Gaminde I, Salazar A, Sánchez A, Ariño J. 2007. A randomized controlled clinical trial of a psychosocial and communication intervention carried out by GPs for patients with medically unexplained symptoms. *Psychol Med.* 37(2):283–294. <https://doi.org/10.1017/S0033291706009536>

Ajam AA, Nguyen XV, Kelly RA, Ladapo JA, Lang EV. 2017. Effects of Interpersonal Skills Training on MRI Operations in a Saturated Market: A Randomized Trial. *J Am Coll Radiol JACR.* 14(7):963–970. <https://doi.org/10.1016/j.jacr.2017.03.015>

Aktas S, Pasinlioğlu T. 2021. The effect of empathy training given to midwives on the empathic communication skills of midwives and the birth satisfaction of mothers giving birth with the help of these midwives: A quasi-experimental study. *J Eval Clin Pract.* 27(4):858–867. <https://doi.org/10.1111/jep.13523>

Alexander C, Sheeler RD, Rasmussen NH, Hayden L. 2015. Teaching an experiential mind-body method to medical students to increase interpersonal skills: a pilot study. *Acad Psychiatry J Am Assoc Dir Psychiatr Resid Train Assoc Acad Psychiatry.* 39(3):316–319. <https://doi.org/10.1007/s40596-014-0159-y>

Alhassan M. 2019. Effect of a 2-day communication skills training on nursing and midwifery students' empathy: a randomised controlled trial. *BMJ Open.* 9(3):e023666. <https://doi.org/10.1136/bmjopen-2018-023666>

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 - Légifrance. [accessed 2022 Feb 25]. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000042320018>

Arrêté du 21 décembre 2021 relatif à l'organisation des épreuves nationales donnant accès au troisième cycle des études de médecine. [place unknown].

Asan O, Montague E. 2014. Using video-based observation research methods in primary care health encounters to evaluate complex interactions. *Inform Prim Care.* 21(4):161–170. <https://doi.org/10.14236/jhi.v21i4.72>

Bank I, Rasenberg EMC, Makkenze-Mangold SH, Nelissen S, van Wijngaarden J, Lovink AG, Rethans J-J. 2021. Fifteen simulated patient working formats to use in communication skills training: Report of a survey. *Med Teach.* 43(12):1391–1397. <https://doi.org/10.1080/0142159X.2021.1948522>

Batt-Rawden SA, Chisolm MS, Anton B, Flickinger TE. 2013. Teaching empathy to medical students: an updated, systematic review. *Acad Med J Assoc Am Med Coll.* 88(8):1171–1177. <https://doi.org/10.1097/ACM.0b013e318299f3e3>

- Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J. 2020. Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills. *PloS One*. 15(4):e0230672. <https://doi.org/10.1371/journal.pone.0230672>
- Bellier A, Labarère J, Putkaradze Z, Cavalie G, Carras S, Pelen F, Paris A, Chaffanjon P. 2022. Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): protocol for a randomised controlled trial. *BMJ Open*. 12(2):e051600. <https://doi.org/10.1136/bmjopen-2021-051600>
- Berger-Estilita J, Fuchs A, Hahn M, Chiang H, Greif R. 2020. Attitudes towards Interprofessional education in the medical curriculum: a systematic review of the literature. *BMC Med Educ*. 20(1):254. <https://doi.org/10.1186/s12909-020-02176-4>
- Blanc G, Bellier A, Di Marco L. 2020. Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en maïeutique. *Sages-Femmes*. 19(4):50–54. <https://doi.org/10.1016/j.sagf.2020.06.013>
- Boissy A, Windover AK, Bokar D, Karafa M, Neuendorf K, Frankel RM, Merlino J, Rothberg MB. 2016. Communication Skills Training for Physicians Improves Patient Satisfaction. *J Gen Intern Med*. 31(7):755–761. <https://doi.org/10.1007/s11606-016-3597-2>
- Boon H, Stewart M. 1998. Patient-physician communication assessment instruments:: 1986 to 1996 in review. *Patient Educ Couns*. 35(3):161–176. [https://doi.org/10.1016/S0738-3991\(98\)00063-9](https://doi.org/10.1016/S0738-3991(98)00063-9)
- Bosméan L, Chaffanjon P, Bellier A. 2022. Impact of physician-patient relationship training on medical students' interpersonal skills during simulated medical consultations: a cross-sectional study. *BMC Med Educ*. 22(1):117. <https://doi.org/10.1186/s12909-022-03171-7>
- Brackett MA, Rivers SE, Shiffman S, Lerner N, Salovey P. 2006. Relating emotional abilities to social functioning: a comparison of self-report and performance measures of emotional intelligence. *J Pers Soc Psychol*. 91(4):780–795. <https://doi.org/10.1037/0022-3514.91.4.780>
- Buckley S, Ambrose L, Anderson E, Coleman JJ, Hensman M, Hirsch C, Hodson J, Morley D, Pittaway S, Stewart J. 2016. Tools for structured team communication in pre-registration health professions education: a Best Evidence Medical Education (BEME) review: BEME Guide No. 41. *Med Teach*. 38(10):966–980. <https://doi.org/10.1080/0142159X.2016.1215412>
- Burt J, Abel G, Elmore N, Campbell J, Roland M, Benson J, Silverman J. 2014. Assessing communication quality of consultations in primary care: initial reliability of the Global Consultation Rating Scale, based on the Calgary-Cambridge Guide to the Medical Interview. *BMJ Open*. 4(3):e004339. <https://doi.org/10.1136/bmjopen-2013-004339>
- Cals JW, Butler CC, Hopstaken RM, Hood K, Dinant G-J. 2009. Effect of point of care testing for C reactive protein and training in communication skills on antibiotic use in lower respiratory tract infections: cluster randomised trial. *BMJ*. 338:b1374. <https://doi.org/10.1136/bmj.b1374>
- Chassany O, Boureau F, Liard F, Bertin P, Serrie A, Ferran P, Keddad K, Jolivet-Landreau I, Marchand S. 2006. Effects of training on general practitioners' management of pain in osteoarthritis: a randomized multicenter study. *J Rheumatol*. 33(9):1827–1834.

- Cheraghi-Sohi S, Bower P. 2008. Can the feedback of patient assessments, brief training, or their combination, improve the interpersonal skills of primary care physicians? A systematic review. *BMC Health Serv Res.* 8:179. <https://doi.org/10.1186/1472-6963-8-179>
- Chessman AW, Blue AV, Gilbert GE, Carey ME, Mainous AG. 2003. Assessing students' communication and interpersonal skills across evaluation settings. *Fam Med.* 35(9):643–648.
- Chichirez C, Purcărea V. 2018. Interpersonal communication in healthcare. *J Med Life.* 11(2):119–122.
- Christian JG, Bessesen DH, Byers TE, Christian KK, Goldstein MG, Bock BC. 2008. Clinic-based support to help overweight patients with type 2 diabetes increase physical activity and lose weight. *Arch Intern Med.* 168(2):141–146. <https://doi.org/10.1001/archinternmed.2007.13>
- Clayton MF, Latimer S, Dunn TW, Haas L. 2011. Assessing patient-centered communication in a family practice setting: how do we measure it, and whose opinion matters? *Patient Educ Couns.* 84(3):294–302. <https://doi.org/10.1016/j.pec.2011.05.027>
- Coleman T. 2000. Using video-recorded consultations for research in primary care: advantages and limitations. *Fam Pract.* 17(5):422–427.
- Cömert M, Zill JM, Christalle E, Dirmaier J, Härter M, Scholl I. 2016. Assessing Communication Skills of Medical Students in Objective Structured Clinical Examinations (OSCE)--A Systematic Review of Rating Scales. *PloS One.* 11(3):e0152717. <https://doi.org/10.1371/journal.pone.0152717>
- Cooper LA, Roter DL, Carson KA, Bone LR, Larson SM, Miller ER, Barr MS, Levine DM. 2011. A randomized trial to improve patient-centered care and hypertension control in underserved primary care patients. *J Gen Intern Med.* 26(11):1297–1304. <https://doi.org/10.1007/s11606-011-1794-6>
- Cuddy MM, Swygert KA, Swanson DB, Jobe AC. 2011. A multilevel analysis of examinee gender, standardized patient gender, and United States medical licensing examination step 2 clinical skills communication and interpersonal skills scores. *Acad Med J Assoc Am Med Coll.* 86(10 Suppl):S17-20. <https://doi.org/10.1097/ACM.0b013e31822a6c05>
- Dennhardt S, Apramian T, Lingard L, Torabi N, Arntfield S. 2016. Rethinking research in the medical humanities: a scoping review and narrative synthesis of quantitative outcome studies. *Med Educ.* 50(3):285–299. <https://doi.org/10.1111/medu.12812>
- Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. 2001. Influence of context effects on health outcomes: a systematic review. *Lancet Lond Engl.* 357(9258):757–762.
- Donnelly MB, Sloan D, Plymale M, Schwartz R. 2000. Assessment of residents' interpersonal skills by faculty proctors and standardized patients: a psychometric analysis. *Acad Med J Assoc Am Med Coll.* 75(10 Suppl):S93-95.
- Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, Lofton S, Wallace M, Goode L, Langdon L, Participants in the American Academy on Physician and Patient's Conference on Education and Evaluation of Competence in Communication and Interpersonal Skills. 2004. Assessing competence in communication and interpersonal skills: the Kalamazoo

II report. *Acad Med J Assoc Am Med Coll.* 79(6):495–507. <https://doi.org/10.1097/00001888-200406000-00002>

Dyche L. 2007. Interpersonal skill in medicine: the essential partner of verbal communication. *J Gen Intern Med.* 22(7):1035–1039. <https://doi.org/10.1007/s11606-007-0153-0>

Dyche L, Swiderski D. 2005. The effect of physician solicitation approaches on ability to identify patient concerns. *J Gen Intern Med.* 20(3):267–270. <https://doi.org/10.1111/j.1525-1497.2005.40266.x>

Egener B, Cole-Kelly K. 2004. Satisfying the patient, but failing the test. *Acad Med J Assoc Am Med Coll.* 79(6):508–510. <https://doi.org/10.1097/00001888-200406000-00003>

Epstein J, Santo RM, Guillemin F. 2015. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol.* 68(4):435–441. <https://doi.org/10.1016/j.jclinepi.2014.11.021>

Fagerlind H, Bergström I, Lindblad AK, Velikova G, Glimelius B, Ring L. 2011. Communication analysis in oncology care. Performance of a combination of a content analysis system and a global scale. *Psychooncology.* 20(9):992–1000. <https://doi.org/10.1002/pon.1808>

Falissard B. 2008. *Mesurer la subjectivité en santé: Perspective méthodologique et statistique.* [place unknown]: Elsevier Masson.

Fassaert T, van Dulmen S, Schellevis F, Bensing J. 2007. Active listening in medical consultations: development of the Active Listening Observation Scale (ALOS-global). *Patient Educ Couns.* 68(3):258–264. <https://doi.org/10.1016/j.pec.2007.06.011>

Fleetwood VA, Veenstra B, Wojtowicz A, Kerchberger J, Velasco J. 2018. Communication through simulation: Developing a curriculum to teach interpersonal skills. *Surgery.* 164(4):802–809. <https://doi.org/10.1016/j.surg.2018.05.037>

Fossli Jensen B, Gulbrandsen P, Benth JS, Dahl FA, Krupat E, Finset A. 2010. Interrater reliability for the Four Habits Coding Scheme as part of a randomized controlled trial. *Patient Educ Couns.* 80(3):405–409. <https://doi.org/10.1016/j.pec.2010.06.032>

Gilligan C, Powell M, Lynagh MC, Ward BM, Lonsdale C, Harvey P, James EL, Rich D, Dewi SP, Nepal S, et al. 2021. Interventions for improving medical students' interpersonal communication in medical consultations. *Cochrane Database Syst Rev.* 2:CD012418. <https://doi.org/10.1002/14651858.CD012418.pub2>

Girgis A, Cockburn J, Butow P, Bowman D, Schofield P, Stojanovski E, D'Este C, Tattersall MHN, Doran C, Turner J. 2009. Improving patient emotional functioning and psychological morbidity: evaluation of a consultation skills training program for oncologists. *Patient Educ Couns.* 77(3):456–462. <https://doi.org/10.1016/j.pec.2009.09.018>

Gulbrandsen P, Jensen BF, Finset A, Blanch-Hartigan D. 2013. Long-term effect of communication training on the relationship between physicians' self-efficacy and performance. *Patient Educ Couns.* 91(2):180–185. <https://doi.org/10.1016/j.pec.2012.11.015>

Ha JF, Longnecker N. 2010. Doctor-patient communication: a review. *Ochsner J.* 10(1):38–43.

- Hall JA, Roter DL. 2002. Do patients talk differently to male and female physicians? A meta-analytic review. *Patient Educ Couns.* 48(3):217–224. [https://doi.org/10.1016/s0738-3991\(02\)00174-x](https://doi.org/10.1016/s0738-3991(02)00174-x)
- Heffron MG, Simson D, Kochar MS. 2007. Competency-based physician education, recertification, and licensure. *WMJ Off Publ State Med Soc Wis.* 106(4):215–218.
- Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687.
- Hulsman RL, Ros WJ, Winnubst JA, Bensing JM. 1999. Teaching clinically experienced physicians communication skills. A review of evaluation studies. *Med Educ.* 33(9):655–668. <https://doi.org/10.1046/j.1365-2923.1999.00519.x>
- Ilcewicz HN, Poirier TI, Pailden J. 2018. Use of mixed-methods approach to assess the impact of a pre-professional health humanities honors course on developing interpersonal skills. *Curr Pharm Teach Learn.* 10(11):1456–1465. <https://doi.org/10.1016/j.cptl.2018.08.010>
- Jefferson L, Bloor K, Birks Y, Hewitt C, Bland M. 2013. Effect of physicians' gender on communication and consultation length: a systematic review and meta-analysis. *J Health Serv Res Policy.* 18(4):242–248. <https://doi.org/10.1177/1355819613486465>
- Jouriles NJ, Emerman CL, Cydulka RK. 2002. Direct observation for assessing emergency medicine core competencies: interpersonal skills. *Acad Emerg Med Off J Soc Acad Emerg Med.* 9(11):1338–1341. <https://doi.org/10.1111/j.1553-2712.2002.tb01598.x>
- Junod Perron N, Sommer J, Louis-Simonet M, Nendaz M. 2015. Teaching communication skills: beyond wishful thinking. *Swiss Med Wkly.* 145:w14064. <https://doi.org/10.4414/smw.2015.14064>
- Kelley JM, Kraft-Todd G, Schapira L, Kossowsky J, Riess H. 2014. The influence of the patient-clinician relationship on healthcare outcomes: a systematic review and meta-analysis of randomized controlled trials. *PloS One.* 9(4):e94207. <https://doi.org/10.1371/journal.pone.0094207>
- Kleinlogel EP, Curdy M, Rodrigues J, Sandi C, Schmid Mast M. 2021. Doppelgänger-based training: Imitating our virtual self to accelerate interpersonal skills learning. *PloS One.* 16(2):e0245960. <https://doi.org/10.1371/journal.pone.0245960>
- Krupat E, Frankel R, Stein T, Irish J. 2006. The Four Habits Coding Scheme: validation of an instrument to assess clinicians' communication behavior. *Patient Educ Couns.* 62(1):38–45. <https://doi.org/10.1016/j.pec.2005.04.015>
- Kurtz S, Silverman J, Draper J. 2004. *Teaching and Learning Communication Skills in Medicine, Second Edition.* 2nd Revised edition edition. Oxford: Radcliffe Publishing Ltd.
- Kurtz SM, Silverman JD. 1996. The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. *Med Educ.* 30(2):83–89.

- Leaune E, Rey-Cadilhac V, Oufker S, Grot S, Strowd R, Rode G, Crandall S. 2021. Medical students attitudes toward and intention to work with the underserved: a systematic review and meta-analysis. *BMC Med Educ.* 21(1):129. <https://doi.org/10.1186/s12909-021-02517-x>
- Lemon TI, Smith RH. 2014. Consultation Content not Consultation Length Improves Patient Satisfaction. *J Fam Med Prim Care.* 3(4):333–339. <https://doi.org/10.4103/2249-4863.148102>
- Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. 1997. Physician-patient communication. The relationship with malpractice claims among primary care physicians and surgeons. *JAMA.* 277(7):553–559.
- Lopes PN, Brackett MA, Nezlek JB, Schütz A, Sellin I, Salovey P. 2004. Emotional intelligence and social interaction. *Pers Soc Psychol Bull.* 30(8):1018–1034. <https://doi.org/10.1177/0146167204264762>
- Makoul G, Krupat E, Chang C-H. 2007. Measuring patient views of physician communication skills: Development and testing of the Communication Assessment Tool. *Patient Education and Counseling.* 67(3):333–342. <https://doi.org/10.1016/j.pec.2007.05.005>
- Mann KV, Ruedy J, Millar N, Andreou P. 2005. Achievement of non-cognitive goals of undergraduate medical education: perceptions of medical students, residents, faculty and other health professionals. *Med Educ.* 39(1):40–48. <https://doi.org/10.1111/j.1365-2929.2004.02031.x>
- Mauksch LB, Dugdale DC, Dodson S, Epstein R. 2008. Relationship, communication, and efficiency in the medical encounter: creating a clinical model from a literature review. *Arch Intern Med.* 168(13):1387–1395. <https://doi.org/10.1001/archinte.168.13.1387>
- Morel A, Hardouin J-B, Pottier P. 2018. Interpersonal Skills Among Fourth- and Fifth-Year Medical Students: Validation of an Assessment Tool. *Simul Healthc J Soc Simul Healthc.* 13(2):88–95. <https://doi.org/10.1097/SIH.0000000000000308>
- Murakami G, Imanaka Y, Kobuse H, Lee J, Goto E. 2010. Patient perceived priorities between technical skills and interpersonal skills: their influence on correlates of patient satisfaction. *J Eval Clin Pract.* 16(3):560–568. <https://doi.org/10.1111/j.1365-2753.2009.01160.x>
- Murden RA, Way DP, Hudson A, Westman JA. 2004. Professionalism deficiencies in a first-quarter doctor-patient relationship course predict poor clinical performance in medical school. *Acad Med J Assoc Am Med Coll.* 79(10 Suppl):S46-48.
- Nelson S, White CF, Hodges BD, Tassone M. 2017. Interprofessional Team Training at the Prelicensure Level: A Review of the Literature. *Acad Med J Assoc Am Med Coll.* 92(5):709–716. <https://doi.org/10.1097/ACM.0000000000001435>
- Neumann M, Edelhäuser F, Tauschel D, Fischer MR, Wirtz M, Woopen C, Haramati A, Scheffer C. 2011. Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Acad Med J Assoc Am Med Coll.* 86(8):996–1009. <https://doi.org/10.1097/ACM.0b013e318221e615>
- Nunnally JC, Bernstein I. 1994. *Psychometric Theory.* 3rd ed. New York: McGraw-Hill.

Olmsted JMD. 1935. The Influence of Claude Bernard on Medicine in the United States and England. *Calif West Med.* 42(2):111–113.

Omura M, Maguire J, Levett-Jones T, Stone TE. 2016. Effectiveness of assertive communication training programs for health professionals and students: a systematic review protocol. *JBI Database Syst Rev Implement Rep.* 14(10):64–71. <https://doi.org/10.11124/JBISRIR-2016-003158>

Perry J, Watkins M, Gilbert A, Rawlinson J. 2013. A systematic review of the evidence on service user involvement in interpersonal skills training of mental health students. *J Psychiatr Ment Health Nurs.* 20(6):525–540. <https://doi.org/10.1111/j.1365-2850.2012.01955.x>

Ram P, Grol R, Rethans JJ, Schouten B, van der Vleuten C, Kester A. 1999. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ.* 33(6):447–454.

Rao JK, Anderson LA, Inui TS, Frankel RM. 2007. Communication Interventions Make A Difference in Conversations Between Physicians and Patients: A Systematic Review of the Evidence. *Med Care.* 45(4):340. <https://doi.org/10.1097/01.mlr.0000254516.04961.d5>

Rees C, Sheard C, McPherson A. 2004. Medical students' views and experiences of methods of teaching and learning communication skills. *Patient Educ Couns.* 54(1):119–121. [https://doi.org/10.1016/S0738-3991\(03\)00196-4](https://doi.org/10.1016/S0738-3991(03)00196-4)

Rees CE, Sheard CE, McPherson AC. 2002. A qualitative study to explore undergraduate medical students' attitudes towards communication skills learning. *Med Teach.* 24(3):289–293. <https://doi.org/10.1080/01421590220134123>

Rider EA, Keefer CH. 2006. Communication skills competencies: definitions and a teaching toolbox. *Med Educ.* 40(7):624–629. <https://doi.org/10.1111/j.1365-2929.2006.02500.x>

Roter DL, Hall JA, Aoki Y. 2002. Physician gender effects in medical communication: a meta-analytic review. *JAMA.* 288(6):756–764.

Ruiz Moral R, Rodríguez Salvador JJ, Pérula De Torres L, Prados Castillejo JA. 2002. [Evolution of the communication profile of family medicine residents]. *Aten Primaria.* 29(3):132–141. [https://doi.org/10.1016/s0212-6567\(02\)70524-6](https://doi.org/10.1016/s0212-6567(02)70524-6)

Sanson-Fisher R, Hobden B, Waller A, Dodd N, Boyd L. 2018. Methodological quality of teaching communication skills to undergraduate medical students: a mapping review. *BMC Med Educ.* 18(1):151. <https://doi.org/10.1186/s12909-018-1265-4>

Schirmer JM, Mauksch L, Lang F, Marvel MK, Zoppi K, Epstein RM, Brock D, Pryzbylski M. 2005. Assessing communication competence: a review of current tools. *Fam Med.* 37(3):184–192.

Scholl I, Nicolai J, Pahlke S, Kriston L, Krupat E, Härter M. 2014. The German version of the Four Habits Coding Scheme - association between physicians' communication and shared decision making skills in the medical encounter. *Patient Educ Couns.* 94(2):224–229. <https://doi.org/10.1016/j.pec.2013.10.006>

- Sequist TD, Fitzmaurice GM, Marshall R, Shaykevich S, Marston A, Safran DG, Ayanian JZ. 2010. Cultural competency training and performance reports to improve diabetes care for black patients: a cluster randomized, controlled trial. *Ann Intern Med.* 152(1):40–46. <https://doi.org/10.7326/0003-4819-152-1-201001050-00009>
- Shelton W, Geppert C, Jankowski J. 2016. The Role of Communication and Interpersonal Skills in Clinical Ethics Consultation: The Need for a Competency in Advanced Ethics Facilitation. *J Clin Ethics.* 27(1):28–38.
- Simpson SA, Shaw C, McNamara R. 2011. What is the most effective way to maintain weight loss in adults? *BMJ.* 343:d8042.
- Stein T, Frankel RM, Krupat E. 2005. Enhancing clinician communication skills in a large healthcare organization: a longitudinal case study. *Patient Educ Couns.* 58(1):4–12. <https://doi.org/10.1016/j.pec.2005.01.014>
- Stepanikova I, Zhang Q, Wieland D, Eleazer GP, Stewart T. 2012. Non-verbal communication between primary care physicians and older patients: how does race matter? *J Gen Intern Med.* 27(5):576–581. <https://doi.org/10.1007/s11606-011-1934-z>
- Stewart J. 1995. Northern practices offer a challenge but few MDs willing to make the move. *CMAJ Can Med Assoc J J Assoc Medicale Can.* 153(5):667–668.
- Stewart M, Brown JB, Donner A, McWhinney IR, Oates J, Weston WW, Jordan J. 2000. The impact of patient-centered care on outcomes. *J Fam Pract.* 49(9):796–804.
- Stubenrauch S, Schneid E-M, Wunsch A, Helmes A, Bertz H, Fritzsche K, Wirsching M, Gözl T. 2012. Development and evaluation of a checklist assessing communication skills of oncologists: the COM-ON-Checklist. *J Eval Clin Pract.* 18(2):225–230. <https://doi.org/10.1111/j.1365-2753.2010.01556.x>
- Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 60(1):34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
- Wang S, Shadrake L, Lyon MJ, Kim H, Yudkowsky R, Hernandez C. 2015. Standardized patient-based assessment of dermatology resident communication and interpersonal skills. *JAMA Dermatol.* 151(3):340–342. <https://doi.org/10.1001/jamadermatol.2014.3646>
- Watters C, Reedy G, Ross A, Morgan NJ, Handslip R, Jaye P. 2015. Does interprofessional simulation increase self-efficacy: a comparative study. *BMJ Open.* 5(1):e005472. <https://doi.org/10.1136/bmjopen-2014-005472>
- Wilson AD, Childs S. 2006. Effects of interventions aimed at changing the length of primary care physicians' consultation. *Cochrane Database Syst Rev.*(1):CD003540. <https://doi.org/10.1002/14651858.CD003540.pub2>
- Wouda JC, van de Wiel HBM. 2012. The communication competency of medical students, residents and consultants. *Patient Educ Couns.* 86(1):57–62. <https://doi.org/10.1016/j.pec.2011.03.011>

Yudkowsky R, Downing SM, Sandlow LJ. 2006. Developing an institution-based assessment of resident communication and interpersonal skills. *Acad Med J Assoc Am Med Coll.* 81(12):1115–1122. <https://doi.org/10.1097/01.ACM.0000246752.00689.bf>

van Zanten M, Boulet JR, McKinley D. 2007. Using standardized patients to assess the interpersonal skills of physicians: six years' experience with a high-stakes certification examination. *Health Commun.* 22(3):195–205. <https://doi.org/10.1080/10410230701626562>

Zick A, Granieri M, Makoul G. 2007. First-year medical students' assessment of their own communication skills: a video-based, open-ended approach. *Patient Educ Couns.* 68(2):161–166. <https://doi.org/10.1016/j.pec.2007.05.018>

Zill JM, Christalle E, Müller E, Härter M, Dirmaier J, Scholl I. 2014. Measurement of physician-patient communication--a systematic review. *PloS One.* 9(12):e112637. <https://doi.org/10.1371/journal.pone.0112637>

Zolnierok KBH, Dimatteo MR. 2009. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care.* 47(8):826–834. <https://doi.org/10.1097/MLR.0b013e31819a5acc>

Zoppi K, Epstein RM. 2002. Is communication a skill? Communication behaviors and being in relation. *Fam Med.* 34(5):319–324.

7. Valorisation des travaux de thèse

- 7.1. **Article 1** : Bellier A, Chaffanjon P, Krupat E, Francois P, Labarère J. Cross-cultural adaptation of the 4-Habits Coding Scheme into French to assess physician communication skills. PLoS ONE. 2020;15(4):e0230672.

Lien : <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0230672>

- 7.2. **Article 2** : Blanc G, Bellier A, Di Marco L. Validation du système de codage des quatre habitudes pour l'évaluation des compétences relationnelles en maïeutique. Sages-Femmes. 2020 Jul 1;19(4):50–4.

Lien : <https://www.em-consulte.com/article/1381124/article/validation-du-systeme-de-codage-des-quatre-habitud>

- 7.3. **Article 3** : Bellier A, Chaffanjon P, Morand P, Palombi O, Francois P, Labarère J. Student characteristics associated with interpersonal skills in medical consultations. BMC Medical Education. 3 mai 2022;22(1):338.

Lien : <https://bmcmmededuc.biomedcentral.com/articles/10.1186/s12909-022-03412-9>

- 7.4. **Article 4** : Bosméan L, Chaffanjon P, Bellier A. Impact of physician-patient relationship training on medical students' interpersonal skills during simulated medical consultations: a cross-sectional study. BMC Med Educ. 2022 Feb 22;22(1):117.

Lien : <https://bmcmmededuc.biomedcentral.com/articles/10.1186/s12909-022-03171-7>

- 7.5. **Article 5** : Bellier A, Labarère J, Putkaradze Z, Cavalie G, Carras S, Pelen F, et al. Effectiveness of a multifaceted intervention to improve interpersonal skills of physicians in medical consultations (EPECREM): protocol for a randomised controlled trial. BMJ Open. 2022 Feb 15;12(2):e051600.

Lien : <https://bmjopen.bmj.com/content/12/2/e051600>

8. Annexes des articles

8.1. Partie 1) Appendix 1 :

		Number of references
#1	"Clinical Competence"[Mesh] OR "Clinical competence"[TW]	92,901
#2	"Communication"[Mesh] OR "Communication"[TW]	530,431
#3	"Relationship competence"[TW]	12
#4	"Attitude of Health Personnel"[Mesh]	156,395
#5	#1 OR #2 OR #3 OR #4	737,156
#6	"Surveys and Questionnaires/standards"[Mesh]	22,779
#7	"Psychometrics"[Mesh] OR "Psychometrics"[TW]	75,879
#8	"Check-list"[TW] OR "Checklist"[TW]	38,779
#9	#6 OR #7 OR #8	129,005
#10	"Physician-Patient Relations"[Mesh]	71,868
#11	"Physician-Patient Interaction"[TW]	321
#12	#10 OR #11	71,976
#13	#5 AND #9 AND #12	528

Appendix 1. Search strategy via PubMed

8.2. Partie 1) Appendix 2 :

Title	Authors	Journal references	Year	Language
Psychometric Properties of the Spanish Adaptation of the Health Care Communication Questionnaire (HCCQ)	C. LEAL COSTA, R. GÓMEZ SÁNCHEZ, S. TIRADO GONZÁLEZ, J. RODRÍGUEZ MARÍN AND C. J. VAN-DER HOFSTADT ROMÁN	Span J Psychol. 18:E96	2015	Spanish
Development and validation of the "FrOCK": Frankfurt observer communication checklist	M. SENNEKAMP, K. GILBERT, F. M. GERLACH AND C. GUETHLIN	Z Evidenz Fortbild Qual Im Gesundheitswesen. 106(8):595-601	2012	German

Appendix 2. Potentially relevant primary studies published in languages other than English and French

8.3. Partie 1) Appendix 3 :

Items - Dimensions	4-HCS	GCRS	ALOS-Global	PCBI	LIV-MAAS	RCS-O	CAT	ECRMP	Verdun Check-list
Initiating the session	X	X	X	X	X
Problem identification		X	X	X
Problem exploration	X	X	X	...	X	...	X	X	X
Patient's perspective	X	X		X	X	X	X	X	
Non-verbal communication	X	X	X	X	X	X	...	X	X
Developing relationship / Empathy	X	X	X	X	X	X	...
Developing communication technique	X	...	X	X	X	X	X	X	X
Providing structure	X	X	X	...	X	X
Providing adequate amount and type of information	X	X	...	X	X	...	X	X	X
Aiding accurate recall and understanding	X	X	...		X	...	X	X	X
Incorporating the patient's perspective	X	X	...	X	X	X	X	X	...
Closure	X	X	X	...	X

Appendix 3. Overlap in item content across standardized instruments

Abbreviations: 4-HCS = 4-Habits Coding Scheme; ALOS-global = Active Listening Observation Scale; CAT = Communication Assessment Tool; ECRMP = "Evaluation des compétences relationnelles du médecin avec le patient"; GCRS = Global Consultation Rating Scale; LIV-MAAS = Liverpool-MAAS; PCBI = Patient-centred Behaviour Coding Instrument; RCS-O = Relational Communication Scale for Observational Measurement

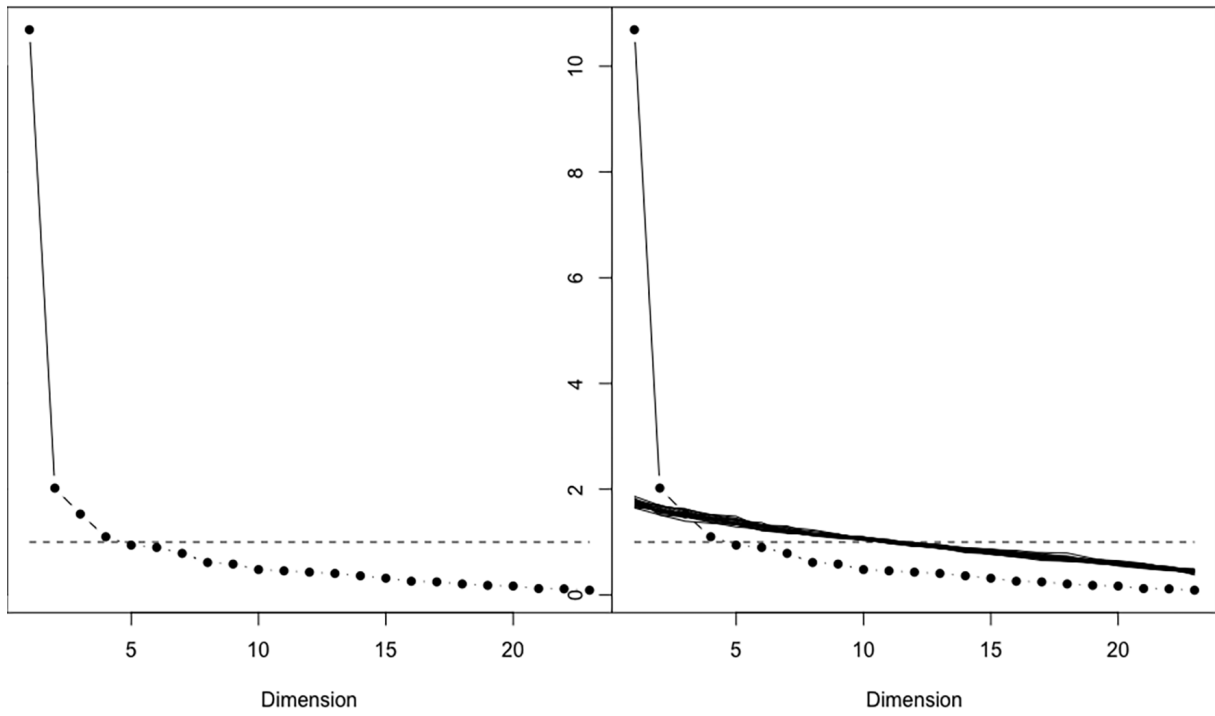
8.4. Partie 2) Appendix 1: Cross-cultural adaptation of the 4-HCS scale into French

Notez chacun des items ci-dessous en utilisant les catégories 1, 3 ou 5. Si vous pensez que la compétence se situe entre ces catégories, vous pouvez utiliser 2 ou 4.

Critères d'évaluation	Degré d'accord
Habitude 1 : S'investir dès le début (/30)	
A. Montre de la familiarité avec le patient	
Le clinicien a besoin de se reporter systématiquement au dossier du patient pour s'y familiariser ou ne fait pas de rapprochement entre la consultation en cours et les antécédents ou le dossier du patient (ou n'a même pas le dossier)	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien fait référence à des consultations précédentes ou aux antécédents du patient, mais ne semble pas bien connaître le sujet	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien connaît bien les antécédents/le dossier médical du patient (par ex., fait référence à des examens récemment pratiqués ou à des informations notées lors de consultations précédentes)	5 <input type="checkbox"/>
B. Accueille les patients chaleureusement	
L'accueil du patient est rapide, impersonnel ou inexistant.	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
L'accueil du clinicien montre qu'il reconnaît le patient, mais l'accueil n'est ni très chaleureux ni personnalisé.	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le patient est accueilli de manière personnelle et chaleureuse (par ex., le clinicien lui demande comment il souhaite qu'on s'adresse à lui, l'appelle par son nom).	5 <input type="checkbox"/>
C. Fait du bavardage	
Le clinicien va droit au but sans aucune tentative de bavardage (ou coupe la parole au patient sèchement et brutalement, ou s'il montre un quelconque intérêt, ce n'est qu'un intérêt superficiel plus tard au cours de la consultation).	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien fait une rapide tentative de bavardage (ne montre pas grand intérêt, met rapidement fin à la conversation avant de poursuivre).	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien fait des commentaires d'ordre non médical afin de mettre le patient à l'aise.	5 <input type="checkbox"/>
D. Utilise principalement des questions ouvertes	
Le clinicien essaie d'identifier le(s) problème(s), principalement à l'aide de questions fermées (conversation saccadée).	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien associe des questions ouvertes et fermées pour tenter d'identifier le(s) problème(s) (par ex. il commence par une question ouverte, puis revient rapidement à des questions fermées).	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien essaie d'identifier le(s) problème(s) principalement à l'aide de questions ouvertes (formule ses questions de façon à ce que le patient raconte sa propre histoire, en essayant de ne pas l'interrompre et avec un minimum de questions fermées).	5 <input type="checkbox"/>
E. Encourage l'expression des inquiétudes du patient	
Le clinicien interrompt le patient et lui coupe la parole lorsqu'il essaie d'approfondir sa réponse (n'est clairement pas très intéressé).	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien n'interrompt pas le patient, mais ne cherche pas à en savoir davantage (écoute, mais n'encourage pas le patient à approfondir ou à développer).	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien encourage le patient à approfondir la conversation et à exprimer ses inquiétudes (par ex., à l'aide de formules incitatives comme « Ah ? », « Dites-m'en plus », « Continuez »).	5 <input type="checkbox"/>
F. Suscite toutes les préoccupations	
Le clinicien s'intéresse immédiatement à la première préoccupation du patient sans tenter de découvrir si le patient a d'autres éventuelles préoccupations.	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien fait référence à d'autres sources d'inquiétudes possibles ou demande rapidement au patient s'il a d'autres inquiétudes, avant de s'intéresser au premier problème évoqué par le patient, ou établit un ordre du jour au fur et à mesure de la consultation.	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien tente de recueillir l'ensemble des préoccupations du patient, en établissant un ordre du jour dès le début de la consultation (le clinicien va bien au-delà de la première préoccupation énoncée).	5 <input type="checkbox"/>
Habitude 2 : Obtenir le point de vue du patient (/15)	
A. Intéressé par la compréhension du problème du patient	
Le clinicien ne cherche pas à comprendre/ne s'intéresse pas au point de vue du patient.	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien ne s'intéresse que brièvement ou superficiellement à la façon dont le patient comprend le problème.	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien fait preuve d'un grand intérêt quant à la façon dont le patient comprend le problème (par ex., il lui demande ce que ses symptômes signifient pour lui).	5 <input type="checkbox"/>
B. Demande les attentes du patient pour la consultation	
Le clinicien n'essaie aucunement de déterminer (ne s'intéresse pas à) ce que le patient attend de la consultation.	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien s'intéresse brièvement à ce que le patient attend de la consultation, mais passe rapidement à la suite.	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien demande (ou réagit avec intérêt à) ce que le patient attend de la consultation (par ex., s'il a des attentes générales ou spécifiques, telles qu'une ordonnance, l'orientation vers un spécialiste).	5 <input type="checkbox"/>
C. Montre de l'intérêt à propos de l'impact sur la vie quotidienne	
Le clinicien n'essaie aucunement de déterminer/ne montre aucun intérêt quant à l'impact que le problème peut avoir sur la vie quotidienne du patient.	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien cherche brièvement à déterminer/ne montre qu'un intérêt limité quant à l'impact que le problème peut avoir sur la vie quotidienne du patient.	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien cherche à déterminer précisément/s'intéresse réellement à l'impact que le problème peut avoir sur la vie quotidienne du patient (au niveau professionnel, familial, sur les activités quotidiennes).	5 <input type="checkbox"/>
Habitude 3 : Faire preuve d'empathie (/20)	
A. Encourage l'expression des émotions	
Le clinicien ne fait preuve d'aucun intérêt quant à l'état émotionnel du patient et/ou décourage ou interrompt le patient quand celui-ci exprime des émotions (signale de façon verbale ou non verbale que le patient ne devrait pas exprimer d'émotions).	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien s'intéresse relativement peu aux émotions du patient, ne l'encourage pas à les exprimer, ou laisse le patient montrer ses émotions, mais l'encourage de façon évidente ou subtile à ne pas s'y attarder.	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien est réceptif aux émotions exprimées par le patient et l'encourage ouvertement (par ex., il incite le patient à poursuivre ou marque des pauses, montre, de façon verbale ou non, qu'il peut exprimer ses émotions).	5 <input type="checkbox"/>
B. Accepte/comprend les sentiments du patient	
Le clinicien n'essaie nullement de montrer qu'il valide les sentiments du patient, voire même les dénigre ou les conteste (par ex., « C'est ridicule de s'inquiéter autant de... »).	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien prend brièvement acte des sentiments du patient, mais ne fait aucun effort pour montrer qu'il reconnaît leur validité/leur bien-fondé.	3 <input type="checkbox"/>
	(4) <input type="checkbox"/>
Le clinicien fait des commentaires qui montrent clairement qu'il reconnaît la validité/le bien-fondé des sentiments du patient (par ex., « Je me sentirais comme vous à votre place », « Je comprends que ça vous puisse inquiéter »).	5 <input type="checkbox"/>
C. Aide à identifier/nommer les sentiments	
Le clinicien n'essaie aucunement d'identifier les sentiments du patient.	1 <input type="checkbox"/>
	(2) <input type="checkbox"/>
Le clinicien fait brièvement référence aux sentiments du patient, mais ne cherche pas vraiment à approfondir en les identifiant ou en les nommant.	3 <input type="checkbox"/>

Le clinicien essaie manifestement d'explorer davantage les sentiments du patient en les identifiant ou en les nommant (par ex., « Qu'est-ce que cela vous fait ? », « Vous me semblez très inquiet au sujet de... »).	(4) 5	<input type="radio"/>
D. Fait preuve d'un comportement non verbal efficace		
Le comportement non verbal du clinicien montre qu'il ne s'intéresse pas, ne se préoccupe pas et/ou marque une certaine distance avec le patient (par ex., peu ou pas de contact du regard, orientation du corps ou utilisation de l'espace inadéquate, ton de voix qui dénote l'ennui).	1	<input type="radio"/>
Le comportement non verbal du clinicien ne témoigne ni réel intérêt ni désintérêt (ou son comportement change au cours de la consultation).	(2) 3	<input type="radio"/>
Le comportement non verbal du clinicien témoigne d'un grand intérêt, il se montre concerné et il établit une certaine proximité avec le patient (par ex., en le regardant dans les yeux, au travers du ton de la voix et de l'orientation du corps) pendant toute la consultation.	(4) 5	<input type="radio"/>
<i>Habitude 4 : S'investir jusqu'à la fin (/50)</i>		
A. Enonce les informations/le diagnostic en fonction des préoccupations du patient		
Le clinicien énonce le diagnostic et les informations en des termes qui correspondent aux références du médecin plutôt que de s'adapter au patient	1	<input type="radio"/>
Le clinicien essaie sommairement d'énoncer le diagnostic et les informations en termes de préoccupations du patient.	(2) 3	<input type="radio"/>
Le clinicien énonce le diagnostic et d'autres informations pertinentes d'une façon qui fait écho aux préoccupations initiales du patient.	(4) 5	<input type="radio"/>
B. Marque des pauses permettant l'intégration de l'information		
Le clinicien donne les informations et poursuit rapidement, sans donner au patient la possibilité de réagir (donnant l'impression que le patient ne se rappellera pas ou n'intégrera pas complètement toutes les informations).	1	<input type="radio"/>
Le clinicien marque une courte pause pour laisser au patient le temps de réagir, puis passe rapidement à autre chose (donnant l'impression que le patient n'a peut-être pas complètement intégré les informations).	(2) 3	<input type="radio"/>
Le clinicien marque une pause après avoir donné les informations au patient afin de lui laisser le temps de réagir et de les intégrer.	(4) 5	<input type="radio"/>
C. Explique clairement/sans jargon médical		
Les informations sont présentées en termes techniques ou inaccessibles au patient (indiquant que le patient n'a probablement pas entièrement ou correctement compris ces informations).	1	<input type="radio"/>
Les informations contiennent du jargon et sont quelque peu difficiles à comprendre	(2) 3	<input type="radio"/>
Les informations sont claires et contiennent peu ou pas de jargon	(4) 5	<input type="radio"/>
D. Explique la nécessité des examens et traitements		
Le clinicien propose de prescrire/préscrit des examens et un traitement, avec peu ou pas d'explication	1	<input type="radio"/>
Le clinicien n'explique que brièvement la raison des examens et du traitement	(2) 3	<input type="radio"/>
Le clinicien explique de façon complète et claire la nécessité des examens et des traitements en cours, passés ou futurs, de façon à ce que le patient comprenne leur importance pour poser le diagnostic et décider du traitement	(4) 5	<input type="radio"/>
E. Vérifie efficacement la compréhension		
Le clinicien ne fait aucun effort pour déterminer si le patient a compris ce qui a été dit	1	<input type="radio"/>
Le clinicien vérifie de façon rapide et insuffisante si le patient a compris	(2) 3	<input type="radio"/>
Le clinicien s'assure que le patient a compris	(4) 5	<input type="radio"/>
F. Encourage la participation à la prise de décision		
Le clinicien ne montre aucun intérêt à impliquer le patient ou décourage manifestement/ignore ses tentatives pour participer aux prises de décision	1	<input type="radio"/>
Le clinicien montre peu d'intérêt à impliquer le patient dans la prise de décision, ou se montre relativement peu enthousiaste lorsque le patient essaie de s'impliquer	(2) 3	<input type="radio"/>
Le clinicien encourage clairement le patient et l'invite à participer activement à la prise de décision	(4) 5	<input type="radio"/>
G. Vérifie l'acceptance du plan thérapeutique		
Le clinicien recommande le traitement sans réellement essayer d'obtenir l'acceptance du patient (volonté ou susceptibilité de suivre le plan thérapeutique).	1	<input type="radio"/>
Le clinicien tente brièvement de déterminer si le patient accepte le plan thérapeutique et passe rapidement à autre chose	(2) 3	<input type="radio"/>
Le clinicien vérifie que le patient accepte le plan thérapeutique, et montre qu'il est prêt à négocier, si besoin	(4) 5	<input type="radio"/>
H. Explore les obstacles à sa mise en oeuvre		
Le clinicien ne mentionne pas les éventuels obstacles à la mise en oeuvre du plan thérapeutique	1	<input type="radio"/>
Le clinicien présente brièvement les potentiels obstacles à la mise en oeuvre du plan thérapeutique.	(2) 3	<input type="radio"/>
Le clinicien présente de façon exhaustive tous les obstacles qui pourraient empêcher la mise en oeuvre du plan thérapeutique	(4) 5	<input type="radio"/>
I. Encourage des questions supplémentaires		
Le clinicien n'encourage aucunement le patient à poser des questions supplémentaires ou les ignore, en grande partie, s'il ne les a pas sollicitées	1	<input type="radio"/>
Le clinicien permet au patient de poser des questions supplémentaires mais ne l'encourage pas à en poser d'autres et y répond sans entrer dans les détails	(2) 3	<input type="radio"/>
Le clinicien encourage ouvertement le patient à poser des questions supplémentaires (et y répond avec une certaine précision).	(4) 5	<input type="radio"/>
J. Etablit un plan de suivi clair		
Le clinicien n'évoque aucun plan de suivi	1	<input type="radio"/>
Le clinicien évoque un suivi, mais ne fait pas de plan précis	(2) 3	<input type="radio"/>
Le clinicien établit un plan de suivi précis et spécifique de la consultation	(4) 5	<input type="radio"/>
TOTAL (/115)		
Ramené à une note /20		

8.5. Partie 2) Appendix 2: Scree plot of the 23 items of the 4-HCS scale



8.6. Partie 5) Appendix 1

Clinical problems	Profiles of portrayals
Significant fatigue with a decrease in physical and mental performance in the context of a depressive syndrome	Reserved and shy patient with a tendency to retain information
History of hemorrhage during dental extraction with genetic background in the context of a coagulation disorder	Disorganized logorrheic patient
New-onset axillary adenopathy after cat scratch in the context of an infectious disease	Anxious and apprehensive patient, taking notes
Episode of tachycardia and moderate exercise-induced asthma in a young athlete	Patient demanding and in a hurry, sometimes aggressive
Incidental finding of microscopic hematuria on urine dipstick with a history of macroscopic hematuria	Curious and knowledgeable patient
Right chest pain with isolated dry cough of recent onset and history of rheumatism in a young woman	Patient with limited understanding
Significant weight loss due to recent-onset diarrhea in an infectious situation.	Minimizing and carefree patient

Table S1. Medical consultation case scenarios and profiles of portrayals

8.7. Partie 5) Appendix 2

Characteristics	Gender		<i>p</i>
	Female (n = 111)	Male (n = 54)	
Research laboratory clerkship, <i>n</i> (%)	17 (15.3)	12 (22.2)	0.028
International clinical placement, <i>n</i> (%)	32 (28.8)	15 (27.8)	0.888
>1 attempts at MCAT, <i>n</i> (%)	49 (44.1)	30 (55.6)	0.170
First year examination score, median (IQR)	74.0 (4.8)	72.7 (4.6)	0.182
Second year examination score, median (IQR)	71.9 (8.6)	68.6 (7.8)	0.008
Third year examination score, median (IQR)	67.6 (10.1)	63.7 (8.0)	0.012
No. medicine clerkships, median (IQR)	3 (1)	3 (1)	0.054
No. Surgery clerkships, median (IQR)	1 (1)	1 (1)	0.259

Table S2. Comparison of undergraduate medical student characteristics according to gender

Abbreviations: MCAT = Medical College Admission Test, IQR = interquartile range (i.e., 25-75th percentiles)

No. items	Invest in the beginning 6	Elicit the patient's perspective 3	Demonstrate empathy 4	Invest in the end 10
Subscale score, median (range)*	61.7 (44.2)	66.7 (63.3)	68.8 (68.8)	67.5 (64.0)
Unadjusted β coefficient (95%CI)				
Male gender	-3.3 (-6.2;-0.4)	-1.7 (-6.0;-2.6)	-5.4 (-9.7;-1.2)	-4.1 (-8.0;-0.2)
Research laboratory placement	2.8 (-0.8;6.4)	4.6 (-0.7;9.9)	4.7 (-0.6;10.0)	5.3 (0.5;10.1)
International clinical placement	5.0 (2.1;8.0)	6.7 (2.3;11.1)	2.3 (-2.1;6.8)	5.2 (1.2;9.2)
Single attempt at MCAT	0.9 (-1.9;3.7)	1.9 (-2.1;6.0)	0.2 (-3.8;4.3)	3.7 (0.1;7.3)
First year examination score	-0.1 (-0.4;0.3)	-0.1 (-0.6;0.5)	0.1 (-0.5;0.6)	0.1 (-0.4;0.6)
Second year examination score	0.0 (-0.2;0.2)	-0.1 (-0.4;0.3)	0.0 (-0.3;0.3)	0.3 (0.0;0.6)
Third year examination score	0.0 (-0.2;0.2)	0.0 (-0.3;0.3)	0.0 (-0.3;0.3)	0.3 (0.0;0.5)
No. medicine clerkships	-1.6 (-3.1;-0.2)	-4.0 (-6.1;-1.9)	-1.3 (-3.5;0.8)	-1.3 (-3.3;0.6)
No. surgery clerkships	-0.4 (-2.1;1.2)	0.7 (-1.7;3.2)	-0.3 (-2.7;2.2)	-1.4 (-3.6;0.8)

* 4-HCS subscale ranged from 0 to 100.

Abbreviations: MCAT = Medical College Admission Test, CI = Confidence Interval; 4-HCS = 4 Habits Coding Scheme

Table S3. Exploratory analysis of univariable associations between 4-HCS subscale scores

8.7. Partie 6) Appendix 1

Variable	Groupe		Habit 1 Median [IQR]	Habit 2 Median [IQR]	Habit 3 Median [IQR]	Habit 4 Median [IQR]	
Gender	General Practitioners	Women	22.5 [20.3, 25.3]	12.0 [9.5, 13.0]	15.5 [14.5, 17.0]	43.0 [40.8, 44.3]	
		Men	21.5 [16.0, 25.0]	10.0 [7.0, 11.5]	13.0 [9.5, 15.5]	35.0 [27.5, 41.5]	
	Hospital Physicians	Women	26.0 [22.0, 28.0]	14.0 [11.0, 15.0]	16.0 [14.0, 18.0]	45.5 [40.3, 49.0]	
		Men	25.0 [22.0, 27.0]	12.0 [10.0, 14.0]	15.0 [14.0, 17.0]	45.0 [40.0, 48.0]	
	Medical Students	Women	21.0 [18.0, 24.0]	12.0 [10.0, 13.2]	14.6 [12.3, 17.0]	37.4 [32.8, 43.0]	
		Men	20.0 [16.9, 23.0]	11.0 [9.0, 13.0]	13.0 [11.0, 16.3]	36.3 [29.4, 42.0]	
	Total	Women	22.0 [18.8, 25.0]*	12.0 [10.0, 14.0]*	15.0 [13.0, 17.0]*	40.0 [34.0, 45.0]*	
		Men	21.0 [17.0, 24.3]	11.0 [9.0, 13.0]	14.0 [11.0, 16.0]	38.5 [30.0, 44.0]	
	Specialty	Hospital Physicians	Surgeon	23.0 [22.0, 27.0]	14.0 [9.5, 14.0]	15.0 [14.0, 16.0]	46.0 [42.0, 49.0]
			Physician	25.0 [22.0, 28.0]	13.0 [11.0, 14.0]	15.5 [14.0, 17.0]	45.0 [40.0, 49.0]
				18.3 [16.3, 20.3]	10.0 [8.5, 11.5]	13.7 [11.8, 15.5]	33.78 [29.3, 37.8]
	Year	Medical Students	2017-18	22.0	13.0	14.0	39.0
			2018-19	[19.0, 24.0]	[11.0, 14.0]	[12.0, 17.0]	[33.0, 45.0]
			2019-20	22.0	12.0	15.0	39.0
2020-21			[18.0, 25.0]	[10.0, 14.0]	[12.0, 18.0]	[32.0, 44.0]	
			24.0 [19.0, 26.0]	12.0 [11.0, 14.0]	16.0 [12.0, 18.0]	40.0 [34.5, 45.0]	

Table S1. Participant's characteristics description by 4-HCS dimensions.

* p-value < 0.001 using a Wilcoxon test

8.8. Partie 6) Appendix 2

Reason for encounter	Code	Group by body system and symptom or diagnoses	Code	Count
Intrauterine contraception	AF03	Family planning	AF	3
Other specified contraception (depot contraception)	AF05			
Examination and encounter for certification purposes	AG03	General and routine examination	AG	3
Routine general health check-up of defined subpopulation	AG04			
Other specified reasons for visit (anesthesia consultation, clinical research)	AR99	Visit for other reasons	AR	22
General pain in multiple sites	AS01	General symptoms, complaints and abnormal findings	AS	6
General weakness or tiredness	AS04			
Other specified abnormal result investigation (elevated creatinine)	AS50			
Complication of medical treatment (post op lymphodema)	AD42	General diagnoses and diseases	AD	2
Lymph gland(s) symptom or complaint	BS01	Symptoms, complaints and abnormal findings of blood, blood-forming organs and immune system	BS	2
Splenomegaly	BS50			
Malignant neopl of blood, blood-forming organs and immune system	BD25	Diagnoses and diseases of blood, blood-forming organs and immune system	BD	24
Benign, uncertain or in-situ neopl of blood, blood-forming organs and immune system	BD26			
Hereditary haemolytic anaemia	BD65			
Other specified and unknown anaemias	BD77			
Coagulation defect	BD78			
Other spec or unknown blood, blood-forming organs, immune system diagnoses or diseases	BD99			
Other specified or unknown malignant digestive neoplasm (pharynx)	DD28			
Eyelid symptoms or complaints	FS09	Symptoms, complaints and abnormal findings of eye	FS	1
Menstrual pain	GS05	Symptoms, complaints and abnormal findings of genital system	GS	5
Irregular or frequent menstruation	GS09			
Concern about breast appearance	GS90			
Genital human papilloma virus infection	GD05	Symptoms, complaints and abnormal findings of genital system	GD	5
Malignant neoplasms of breast	GD27			
Cervical disease	GD65			
Other spec and unknown diagnoses and diseases of genital system (Ovarian cyst)	GD99			
Infection of circulatory system	KD01			
Atrial fibrillation or flutter	KD68			
Hypertension, uncomplicated	KD73			
Hypertension, complicated	KD74			
Atherosclerosis or peripheral vascular disease	KD76			
Other spec and unknown diagnoses and diseases of the circulatory system	KD99			
Chronic widespread pain	LS18	Symptoms, complaints and abnormal findings of musculoskeletal system	LS	1
Fracture of radius or ulna or both	LD35	Diagnoses and diseases of musculoskeletal system	LD	8
Fracture of femur	LD38			
Other specified and unknown fracture (pelvis, vertebral column)	LD39			
Dislocation or subluxation (shoulder)	LD48			
Tennis elbow	LD73			
Other spec diagnoses and diseases of musculoskeletal system	LD99			
Cerebrovascular disease	ND70			
Carpal tunnel syndrome	ND76			
Nose symptoms or complaints	RS10	Symptoms, complaints and abnormal findings of respiratory system	RS	2
Chronic ulcer of skin	SD77	Diagnoses and diseases of skin	SD	2
Other spec or unknown diagnoses and diseases of skin (Erythema nodosum)	SD99			
Type 2 diabetes mellitus	TD72	Diagnoses and diseases of endocrine, metabolic and nutritional system	TD	3
Other spec or unknown endocrine, metabolic, nutritional diagnoses and diseases	TD99			
Nephrosis	UD65	Symptoms, complaints and abnormal findings of urinary system	UD	4
Chronic kidney disease	UD66			

Table S2. Reasons for encounter using ICPC-3 at the body system

8.9. Partie 6) Appendix 3

Reason for encounter	Code	Group by 3-category system	Code	Count
Intrauterine contraception	AF03	General consultation for systemic or preventive reason	G	28
Other specified contraception (depot contraception)	AF05			
Examination and encounter for certification purposes	AG03			
Routine general health check-up of defined subpopulation	AG04			
Other specified reasons for visit (anesthesia consultation, clinical research)	AR99			
Complication of medical treatment (post op lymphodema)	AD42	Consultation following a diagnostic	D	66
Malignant neopl of blood, blood-forming organs and immune system	BD25			
Benign, uncertain or in-situ neopl of blood, blood-forming organs and immune system	BD26			
Hereditary haemolytic anaemia	BD65			
Other specified and unknown anaemias	BD77			
Coagulation defect	BD78			
Other spec or unknown blood, blood-forming organs, immune system diagnoses or diseases	BD99			
Other specified or unknown malignant digestive neoplasm (pharynx)	DD28			
Genital human papilloma virus infection	GD05			
Malignant neoplasms of breast	GD27			
Cervical disease	GD65			
Other spec and unknown diagnoses and diseases of genital system (Ovarian cyst)	GD99			
Infection of circulatory system	KD01			
Atrial fibrillation or flutter	KD68			
Hypertension, uncomplicated	KD73			
Hypertension, complicated	KD74			
Atherosclerosis or peripheral vascular disease	KD76			
Other spec and unknown diagnoses and diseases of the circulatory system	KD99			
Fracture of radius or ulna or both	LD35			
Fracture of femur	LD38			
Other specified and unknown fracture (pelvis, vertebral column)	LD39			
Dislocation or subluxation (shoulder)	LD48			
Tennis elbow	LD73			
Other spec diagnoses and diseases of musculoskeletal system	LD99			
Cerebrovascular disease	ND70			
Carpal tunnel syndrome	ND76			
Chronic ulcer of skin	SD77			
Other spec or unknown diagnoses and diseases of skin (Erythema nodosum)	SD99			
Type 2 diabetes mellitus	TD72			
Other spec or unknown endocrine, metabolic, nutritional diagnoses and diseases	TD99			
Nephrosis	UD65			
Chronic kidney disease	UD66			
General pain in multiple sites	AS01			
General weakness or tiredness	AS04			
Other specified abnormal result investigation (elevated creatinine)	AS50			
Eyelid symptoms or complaints	FS09			
Menstrual pain	GS05			
Irregular or frequent menstruation	GS09			
Concern about breast appearance	GS90			
Chronic widespread pain	LS18			
Nose symptoms or complaints	RS10			

Table S3. Reason for encounter using the ICPC-3 at the 3-category system level