

THESE DE DOCTORAT DE L'UNIVERSITE PARIS DIDEROT (PARIS 7)

Ecole doctorale 393 Pierre Louis de santé publique

Santé publique : Epidémiologie et Sciences de l'Information Biomédicale

Spécialité : Epidémiologie Clinique

Présenté par

Viet-Thi TRAN

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE PARIS DIDEROT – PARIS VII

Développement d'outils de mesure du fardeau du traitement chez des patients souffrant de plusieurs maladies chroniques

Thèse dirigée par le Pr Philippe RAVAUD

Présentée et soutenue publiquement

Le 28 avril 2015

Devant le jury composé de :

Madame le Professeur Véronique Rivain	Rapporteur
Madame le Professeur Marie Zins	Rapporteur
Madame le Professeur France Mentré	Examineur
Monsieur le Professeur Jean François Bergmann	Examineur
Monsieur le Professeur Serge Gilberg	Examineur
Monsieur le Professeur Philippe Ravaud	Directeur de thèse

Remerciements

En tout premier lieu, je tiens à remercier Philippe Ravaud. Il m'a fait bénéficier durant ces trois années de ses conseils avisés, de son jugement abrupt mais souvent sans faille et de son grand savoir méthodologique. Je lui suis reconnaissant d'avoir été disponible en toutes circonstances. Je suis très heureux qu'il ait accepté de guider mes premiers pas en recherche, et qu'il m'ait permis de travailler avec lui sur des sujets passionnants

Je remercie le Professeur Véronique Sébille et le Docteur Marie Zins pour m'avoir accordé l'immense honneur de rapporter ma thèse et pour l'intérêt qu'ils ont lui porté.

Merci également aux Professeurs Jean-François Bergmann, France Mentre et Serge Gilberg de participer à mon jury de thèse.

J'accorde une pensée particulière au Pr Bruno Falissard, avec qui j'espère continuer à travailler à l'avenir et au Pr Victor Montori pour m'avoir aidé et conseillé dans mon travail.

Je suis reconnaissant à l'ensemble de l'équipe INSERM U1153 qui a supporté le fardeau du fardeau du traitement.

J'accorde en particulier une pensée à Raphaël Porcher qui a répondu à toutes mes interrogations statistiques... même si je n'ai souvent pas compris les arcanes de son art. Je remercie (en énumérant (l'ordre n'est pas par importance décroissante)) : 1) Aïda qui a été un cobaye du processus de thèse avant moi ; 2) Aurélie qui m'a nourri de chocolats et pour son titanesque travail de relecture; 3) Isabelle P. qui m'a permis de téléphoner dans son bureau ; 4) Youri qui m'a appris la valeur du GOLD et du persil ; 5) Amélie et Perrine qui me permettent de ne pas prendre mon café seul ; 6) Anne et Céline qui sont là de temps en temps ; 7) Stéphanie S. qui écoute toujours mes idées foireuses avec patience ; 8) Caroline qui a été ma relectrice anglaise officielle et officieuse ; 9) Audrey et Magalie qui m'aident à connaître tous les ragots du service ; 10) Agnès pour sa bonne humeur au quotidien; 11) Ludovic pour son amitié silencieuse ; 12) Elise pour ses beaux dessins ; 13) Alexandre qui aime les arcs-en ciel ; et 14) les stagiaires que je ne connais toujours pas...

Je remercie aussi Michel Nougairède pour m'avoir supporté en stage et pour m'aider depuis dans mes différents projets, les chefs de clinique et maître de conférences du DMG et tous les autres membres du DMG aussi. Je remercie également Clémence Cambon, Giao Do-Pham et mes nombreux co-internes qui m'ont supporté durant toutes ces années.

Merci à Ba et Me et Chi parce que c'est ma famille, Sabine-la-congolaise parce que j'espère qu'elle le sera aussi un jour, Jeremy, Claire et Camille parce qu'ils sont beaux et cools, et aux très nombreux autres que je n'ai pas cités, mais à qui je pense très fort (ou pas). Leur présence a été très importante durant mes années d'étude et est très importante dans ma vie.

PS : toute ressemblance avec des remerciements déjà lus n'est pas totalement fortuite.

Résumé

Aujourd'hui, la prise en charge des patients multimorbides repose sur des stratégies thérapeutiques complexes associant de multiples traitements, consultations, examens et changements d'habitudes de vie. L'ensemble de ces stratégies peut alors représenter un fardeau du traitement pour le patient. Le fardeau du traitement est associé à l'observance thérapeutique et pourrait donc contribuer à la morbidité des patients. L'objectif de cette thèse était de développer des outils pour évaluer le fardeau du traitement.

Nous avons développé le premier outil permettant d'évaluer le fardeau du traitement chez des patients souffrant de maladies chroniques, le TBQ. La validité et la fiabilité de cet instrument a été évaluée en France, auprès de 502 patients ambulatoires et hospitalisés.

Dans un second temps, nous avons adapté le TBQ en anglais par une méthode de traduction-rétro-translation. Les propriétés de l'instrument ont été ensuite évaluées à l'aide d'une plateforme internet auprès de 610 patients vivant dans des pays anglophones.

Le fardeau du traitement est étroitement lié au contexte du patient, en particulier au système de soins auquel il appartient. Afin de développer des instruments permettant de comparer le fardeau entre différents pays et contextes de soins, nous avons dressé une liste exhaustive des composantes du fardeau du traitement au cours d'une étude qualitative. Cette étude, réalisée à l'aide d'un outil internet, a impliqué 1053 patients vivant dans 34 pays différents. Elle nous a permis de créer une taxonomie du fardeau du traitement utilisable quelque soit le pays, le traitement ou la maladie du patient.

Nous avons saisi l'opportunité de cette étude qualitative de grande taille pour travailler sur le concept de saturation des données. Nous avons réalisé des simulations de Monte Carlo pour réarranger les données obtenues précédemment afin d'estimer le point de saturation des données. Ce travail nous a permis de comparer l'efficacité de différentes stratégies de recrutement dans les études qualitatives, décrites dans la littérature.

Nos travaux ont permis de définir le concept de fardeau du traitement et de développer les premiers outils pour le mesurer. Les prochaines étapes consisteront à développer et évaluer des programmes thérapeutiques efficaces minimisant ce fardeau.

Abstract

As chronic diseases are becoming more prevalent, their management strategies are becoming more complicated. Patients are feeling increasingly burdened by their treatments. This burden of treatment represents the “work of being a patient”. It affects treatment adherence, quality of life and outcomes. In order to take into account this burden, clinicians and researchers needed tools to assess it.

Our first project was the development of the first tool to assess the burden of treatment for patients with chronic conditions: the Treatment Burden Questionnaire (TBQ). Validity and reliability of the tool was tested with 502 in- and outpatients, in France.

Our second project was the adaptation of the TBQ in English by a forward-backward translation method. Measurement properties of the adapted instrument were assessed online and involved 610 patients with chronic conditions from the United States, United Kingdom, Canada, Australia and New Zealand.

The burden of treatment is associated with patients’ context. Patients with similar conditions and treatment regimens may have different burdens depending on their education, culture, beliefs, social support, financial capacities, formal and informal support resources, and healthcare context. Thus, in our third project, we performed a multi-country qualitative study using an online survey to describe and classify the components of the burden of treatment across multiple countries and settings. This study involved 1,053 patients, from 34 different countries. Our results may inform the development of cross cultural instruments to measure the burden of treatment, allowing comparison between countries and contexts.

Our fourth project took the opportunity of this large Web-based qualitative study to assess how different sampling strategies approached data saturation by resampling the data generated. To our knowledge, this was the first study to use resampling techniques for investigating a statistical demonstration of saturation within a dataset. Our results might help researchers predict sample size in future qualitative research.

With our work, we developed the first tools to assess the burden of treatment. Further studies should now focus on interventions intended to mitigate the burden of treatment to develop a Minimally Disruptive Medicine.

Laboratoire d'accueil

Centre de recherche Epidémiologie et Statistique INSERM U1153,

Équipe : Méthodes en évaluation thérapeutique des maladies chroniques

Centre d'Epidémiologie Clinique Hôpital Hôtel-Dieu

1, place du parvis Notre-Dame

75181 Paris, France

Publications liées à la thèse

Ce travail de doctorat s'appuie sur les articles suivants :

Tran VT, Montori VM, Eton DT, Baruch D, Falissard B, Ravaud P. Development and description of measurement properties of an instrument to assess Treatment Burden among patients with multiple chronic conditions. *BMC Medicine*. Jul 4 2012; 10(1):68.

Tran VT, Harrington M, Montori VM, Barnes C, Wicks P, Ravaud P. Adaptation and validation of the Treatment Burden Questionnaire (TBQ) in English using an internet platform. *BMC Medicine*. 2014; 12:109.

Tran VT, Barnes C, Montori VM, Falissard B, Ravaud P. Taxonomy of the burden of treatment: a multicountry Web-based qualitative study of patients with chronic conditions. *Soumis à BMC Medicine*. 2015.

Tran VT, Porcher R, Falissard B, Ravaud P. Determining sample size in Web-based qualitative studies: an example of a multinational study on the burden of treatment. *In Soumis au Journal of Clinical Epidemiology*. 2015.

Table des matières

1.	Introduction	13
1.1.	Maladies chroniques et multimorbidité	14
1.2.	Prise en charge actuelle des maladies chroniques.....	18
1.3.	Fardeau du traitement	25
2.	Objectif de la thèse.....	30
3.	Mesurer le fardeau du traitement	32
3.1.	Mesurer la santé et évaluations rapportées par le patient	32
3.2.	Théorie classique des tests.....	33
3.3.	Validité et fiabilité.....	35
3.4.	Importance de l'évaluation de la validité et de la fiabilité	37
3.5.	Critères de qualité des instruments de mesures provenant du patient	38
3.6.	Mesures existantes de la satisfaction du patient vis-à-vis de son traitement ...	43
3.7.	Développement d'un instrument de mesure du fardeau du traitement	48
3.7.1.	Méthodes	48
3.7.2.	Résultats	51
3.7.3.	Discussion.....	52
3.7.4.	Article	56
4.	Adapter la mesure du fardeau du traitement en Anglais	81
4.1.	Adaptation linguistique et transculturelle.....	81
4.2.	Recommandations pour l'adaptation transculturelle des mesures en santé	83
4.3.	Adaptation et validation de notre outil de mesure du fardeau du traitement en Anglais	85
4.3.1.	Méthodes	85
4.3.2.	Résultats	86
4.3.3.	Discussion.....	87
4.3.4.	Article	91
5.	Décrire les composantes, facteurs aggravants et conséquences du fardeau du traitement	108
5.1.	Etudes qualitatives sur le fardeau du traitement	109
5.2.	Décrire le fardeau du traitement dans plusieurs pays et contextes de soins différents	113

5.2.1.	Méthodes	114
5.2.2.	Résultats	116
5.2.3.	Discussion.....	118
5.2.4.	Article	122
6.	Taille d'échantillon dans les études qualitatives	162
6.1.	Définition de la saturation des données	162
6.2.	Méthodes d'estimation du point de saturation des données.....	163
6.3.	Estimer le point de saturation des données dans les études qualitative en ligne 166	
6.3.1.	Méthode.....	167
6.3.2.	Résultats	168
6.3.3.	Discussion.....	170
6.3.4.	Article	172
7.	Implications.....	187
7.1.	Principaux résultats.....	187
7.2.	Implications méthodologiques : repenser le développement d'instruments de mesure provenant du patient.....	187
7.2.1.	Améliorer les méthodes actuelles d'adaptation transculturelle	187
7.2.2.	Etudes utilisant des questionnaires en ligne : le futur ?	189
7.3.	Implications pour les patients : repenser la prise en charge des patients avec des maladies chroniques	191
7.3.1.	Développer des recommandations adaptées à la multimorbidité.....	191
7.3.2.	Lutter contre les prescriptions inutiles	192
7.3.3.	Dépasser les recommandations pour individualiser la médecine	193
7.3.4.	Changer le modèle de consultation.....	194
7.3.5.	Adapter la consultation à la complexité	196
7.3.6.	Adapter la typologie des maladies chroniques à leurs prises en charge ..	197
7.3.7.	Améliorer la coordination entre soignants.....	198
7.3.8.	S'adapter au patient du futur.....	199
8.	Perspectives	202
8.1.	Perspectives méthodologiques	202
8.1.1.	Comparer les méthodes de recueil de données dans les études qualitatives 202	

8.1.2.	Améliorer l'évaluation de la validité de mesures provenant du patient ..	202
8.1.3.	Estimer le nombre de sujets nécessaires dans les études qualitatives	203
8.2.	Perspectives de recherches sur le thème du fardeau du traitement.....	204
8.2.1.	Evaluer la part du fardeau du traitement portée par le patient et ses aidants	204
8.2.2.	Identifier ce que les patients aimeraient modifier dans leur prise en charge	205
8.2.3.	Evaluer l'impact de consultations dédiées à la multimorbidité sur la morbidité et la qualité de vie des patients chroniques.	205
9.	Conclusion	207
10.	Bibliographie	209

1. Introduction

Le fardeau du traitement représente les contraintes liées à tout ce que fait un patient pour se soigner: prendre des médicaments, faire des examens, consulter différents médecins, changer ses habitudes de vie... Ce fardeau du traitement est un phénomène nouveau [1], né des changements du monde médical au cours des dernières décennies. Tout d'abord, la prévalence des maladies chroniques augmente et de plus en plus de patients sont multimorbides. Ensuite, l'amélioration des techniques médicales conduit à proposer aux patients des programmes thérapeutiques aux modalités de suivi de plus en plus complexes. Enfin, de par la spécialisation des soins, les patients multimorbides sont souvent pris en charge par de multiples soignants, chacun s'attachant à traiter un aspect particulier de l'individu.

A l'heure actuelle, il existe peu d'information sur l'impact des interactions entre toutes les interventions proposées au patient, et notamment leur accumulation. Nous sommes donc face à une situation paradoxale où il est possible que ce que l'on demande à un patient pour se soigner, nuise parfois plus à sa santé qu'il ne l'améliore.

1.1. Maladies chroniques et multimorbidité

Les progrès de la médecine ont changé le paysage médical moderne. Par exemple, avant 1920, le diabète de type I était une maladie mortelle. Lorsqu'un patient était diagnostiqué, il décédait en quelques semaines. La découverte du rôle de l'insuline a transformé cette maladie autrefois aiguë et fatale en maladie chronique [2]. La maladie coronaire était une maladie aiguë avant les années 1960. De même pour l'infection par le virus du VIH. Les progrès de la médecine permettent aux patients de vivre plus longtemps avec une maladie chronique. Nous observons donc une augmentation de leur prévalence. Une étude observationnelle transversale réalisée à partir de 314 cabinets de médecine générale en Ecosse, représentant 1,751,841 patients, a mis en évidence que 42.2% (IC 95% [42.1 - 42.3]) des patients avaient au moins une maladie chronique [3]. Et, « Mère Nature étant sans merci, avoir une maladie n'immunise que rarement contre le fait d'en avoir une autre » [4]. Ainsi, les patients ont souvent plusieurs maladies chroniques.

La multimorbidité est définie comme « le fait qu'un patient soit affecté par deux ou plus maladies chroniques » [5]. Dans l'étude précédemment citée, 23.2% (IC 95% [23.1 - 23.2]) des patients avaient plusieurs (≥ 2) maladies chroniques. D'autres études trouvaient des résultats sensiblement différents concernant la prévalence de la multimorbidité [6-10] ; ces différences pouvant venir de la difficulté à dénombrer le nombre de maladies d'un patient donné [11] et des définitions employées (par exemple, dans la définition de l'European General Practice Network (EGPRN), la multimorbidité prend en compte en plus des maladies chroniques, les problèmes psychosociaux et les facteurs de risque somatiques).

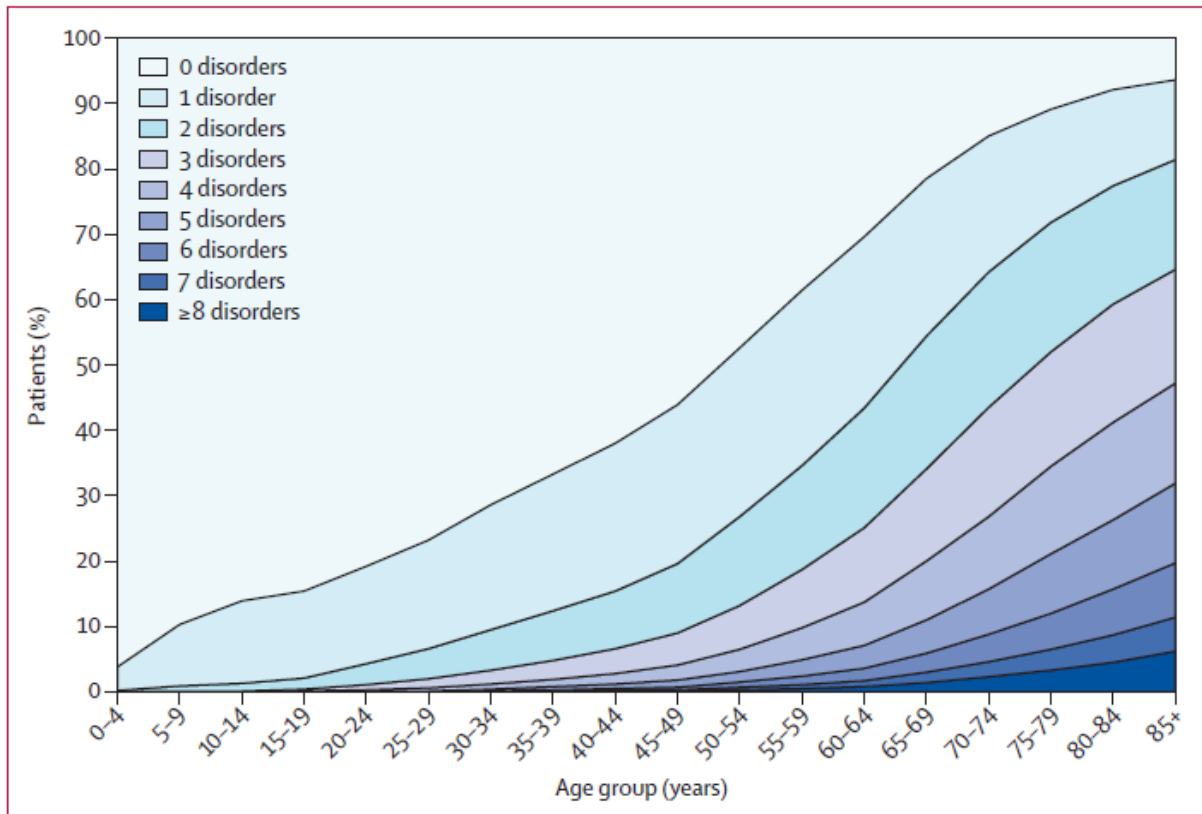
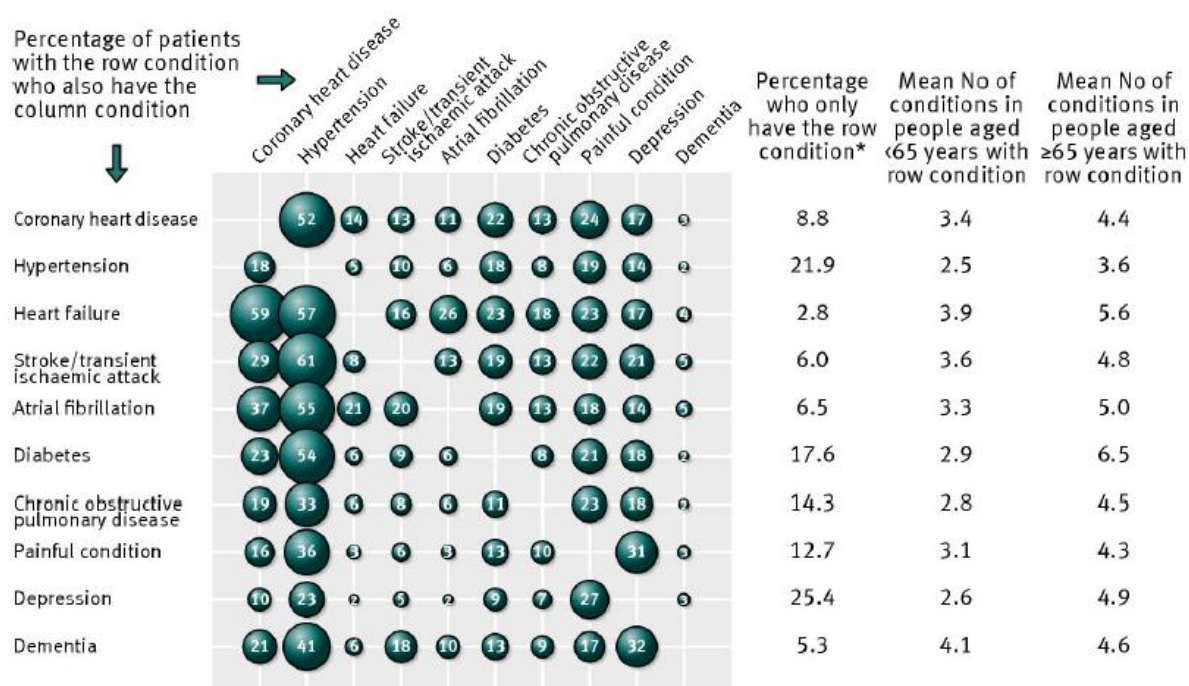


Figure 1 : Nombre de maladies chroniques selon l'âge du patient. Adapté de la source [3]

Le principal facteur de risque d'être multimorbide est l'âge: 50% des patients de plus de 50 ans ont au moins une maladie chronique et la plupart des patients de plus de 65 ans sont multimorbides (**Figure 1**). Cependant, d'autres facteurs peuvent jouer : 1) le niveau socioéconomique (la prévalence de la multimorbidité chez les adultes vivant dans des zones défavorisées économiquement est égale à celle de patients de 10 à 15 ans plus âgés, vivant dans des zones favorisées économiquement [3]) ; 2) les habitudes de vie (tabagisme, alcool, etc.) et/ou 3) les facteurs personnels (obésité, etc.) [12].

Les associations entre maladies ne sont pas aléatoires: certaines maladies sont associées à des facteurs de risque communs, d'autres sont les complications de maladies initiales [13]. Une revue systématique de 14 études cherchant l'association non-aléatoire de maladies chroniques a mis en évidence 3 groupes de maladies

associées: 1) les maladies métaboliques, cardio-pulmonaires et cardiovasculaires ; 2) les troubles de l'humeur, les troubles anxieux et psychosomatiques; et 3) les maladies ostéo-articulaires [14]. Ces résultats concordent avec ceux l'étude de Barnett et al. (Figure 2) où l'on retrouve à la fois des associations entre des maladies « proches » physiologiquement se compliquant les unes les autres (maladie coronaire, insuffisance cardiaque, hypertension artérielle, etc.) ; et des associations entre des maladies ayant des facteurs de risque communs (par exemple : la démence complique rarement une autre maladie, mais la plupart des patients déments présentent des maladies « physiques ») [15].



* Percentage who do not have one of 39 other conditions in the full count

Comorbidity of 10 common conditions among UK primary care patients²

Figure 2: Associations entre maladies chroniques. Adapté de la source [15]

L'augmentation de la prévalence des maladies chroniques est également liée à notre capacité à les diagnostiquer. Avec l'évolution technique, il est possible de détecter des anomalies de plus en plus tôt, souvent à des stades asymptomatiques. Ces anomalies posent la question du seuil entre normal et pathologique. Sous la pression de l'industrie, ces seuils sont régulièrement abaissés, entraînant l'apparition de « maladies » (pré diabète, pré HTA...) dont le caractère pathologique est discutable. Par exemple, la modification de la définition ostéo-densitométrique de l'ostéoporose de 0.5 points augmenterait le nombre de patients considérés comme malades de 6,780,600, aux Etats Unis [16].

Le fait d'avoir une maladie chronique, et d'autant plus d'être multimorbide, joue un impact négatif sur la santé des patients et leur qualité de vie. Ainsi, sur les 57 millions de décès qui se sont produits dans le monde en 2008, 36 millions, soit près des deux tiers, ont été dus aux maladies chroniques et principalement aux maladies cardiovasculaires, aux cancers, au diabète et aux pneumopathies chroniques [17]. Outre la mortalité, la multimorbidité a également un impact sur les taux d'hospitalisation. Une étude observationnelle auprès de 1,217,103 patients de plus de 65 ans, aux Etats-Unis, a montré que la proportion d'hospitalisations liées à une prise en charge ambulatoire inadéquate, augmentait en fonction du nombre de maladies chroniques du patient (OR= 7.5, IC95% [6.5-8.6] et OR=98.5 IC95% [86.1-112.7] pour les patients avec respectivement 1 et 4 maladies chroniques comparé à ceux n'ayant pas de maladies chroniques) [9]. De manière évidente, un patient multimorbide est en moins bon état de santé –plus fragile- qu'un patient « monomorbide », mais cela ne suffit pas à expliquer cette augmentation du risque d'hospitalisation chez ces patients. Plusieurs pistes existent : 1) lorsqu'un patient souffre d'une -ou plusieurs- maladie chronique, ses

autres problèmes de santé sont souvent sous diagnostiqués ou sous traités [4, 18]; 2) le nombre de traitements et le nombre d'interactions médicamenteuses augmentent avec le nombre de maladies [19] avec un risque iatrogène accru; et 3) il est possible, du fait de l'accumulation des interventions, que les patients multimorbides n'arrivent pas à être observants pour l'ensemble de leurs traitements et arrêtent des thérapeutiques utiles.

Devant l'accroissement de la prévalence des maladies chroniques, il est donc nécessaire de réfléchir aux raisons et solutions pour mieux appréhender la prise en charge des patients avec des maladies chroniques.

1.2. Prise en charge actuelle des maladies chroniques

Comme évoqué précédemment, les maladies chroniques sont un phénomène relativement nouveau dans le monde médical. Jusqu'au début du siècle, la plupart des consultations médicales étaient relatives à des problèmes aigus où l'enjeu était de diagnostiquer la maladie afin de prescrire le meilleur traitement. Ceci a forgé une médecine centrée sur la maladie, que ce soit : 1) lors de la formation des médecins, qui repose principalement sur l'apprentissage du diagnostic et le traitement de maladies spécifiques; 2) dans la prise en charge des patients, fragmentée entre plusieurs médecins ou centres, chacun s'occupant d'un problème de santé donné [20]; ou 3) dans l'utilisation de recommandations de pratique clinique, établies maladie par maladie.

La formation des médecins est centrée sur la maladie [21] et met l'accent sur l'exhaustivité. Par exemple, lorsqu'ils examinent des patients, les étudiants sont encouragés et félicités lorsqu'ils énumèrent tous les diagnostics possibles. Ceci est à

l'origine d'une croyance où l'étudiant qui connaît tous les diagnostics est le meilleur [22]. Ainsi, les médecins sont encouragés à prescrire des examens pour éliminer les diagnostics différentiels de chaque symptôme et, une fois le diagnostic établi, à traiter et prodiguer de multiples conseils, de manière exhaustive, parfois sans réfléchir aux conséquences pour le patient.

De plus en plus de soins et pratiques médicales nécessitent un haut niveau de technicité et de spécialisation. Pour assurer la meilleure prise en charge possible de patients complexes et multimorbides, il est donc nécessaire de faire intervenir plusieurs soignants. Cependant, des soins fragmentés entre plusieurs spécialistes impliquent une coordination optimale entre les différents intervenants. Malheureusement, cette coordination est souvent imparfaite. Dans une étude transversale portant sur 253 médecins et 1614 patients aux Etats-Unis, les médecins déclaraient manquer d'au moins une information importante lors de leurs consultations. Ces informations pouvaient être, par exemple, des résultats d'examens, des comptes-rendus de consultations ou d'hospitalisation ou des ordonnances de traitements. Les médecins considéraient que ces informations inaccessibles avaient un effet négatif sur la santé des patients dans 44% des cas et pouvaient retarder la prise en charge des patients dans 59.5% des cas [23]. Lorsqu'on examinait plus spécifiquement les relations entre ville et hôpital, seuls 17-34% des patients avaient un compte-rendu hospitalier lors de leur première consultation ambulatoire post-hospitalisation (51-77% à 4 semaines) et, pour ceux qui avaient un compte-rendu, les médecins estimaient qu'une information était manquante dans au moins 33% des cas [24]. La fragmentation des soins ne concerne pas uniquement l'absence de compte-rendu, mais également les problèmes de communication « directe » entre les intervenants, ce qui

est à l'origine de difficultés pour arrêter ou modifier les prescriptions d'autres médecins [25] (Figure 3).

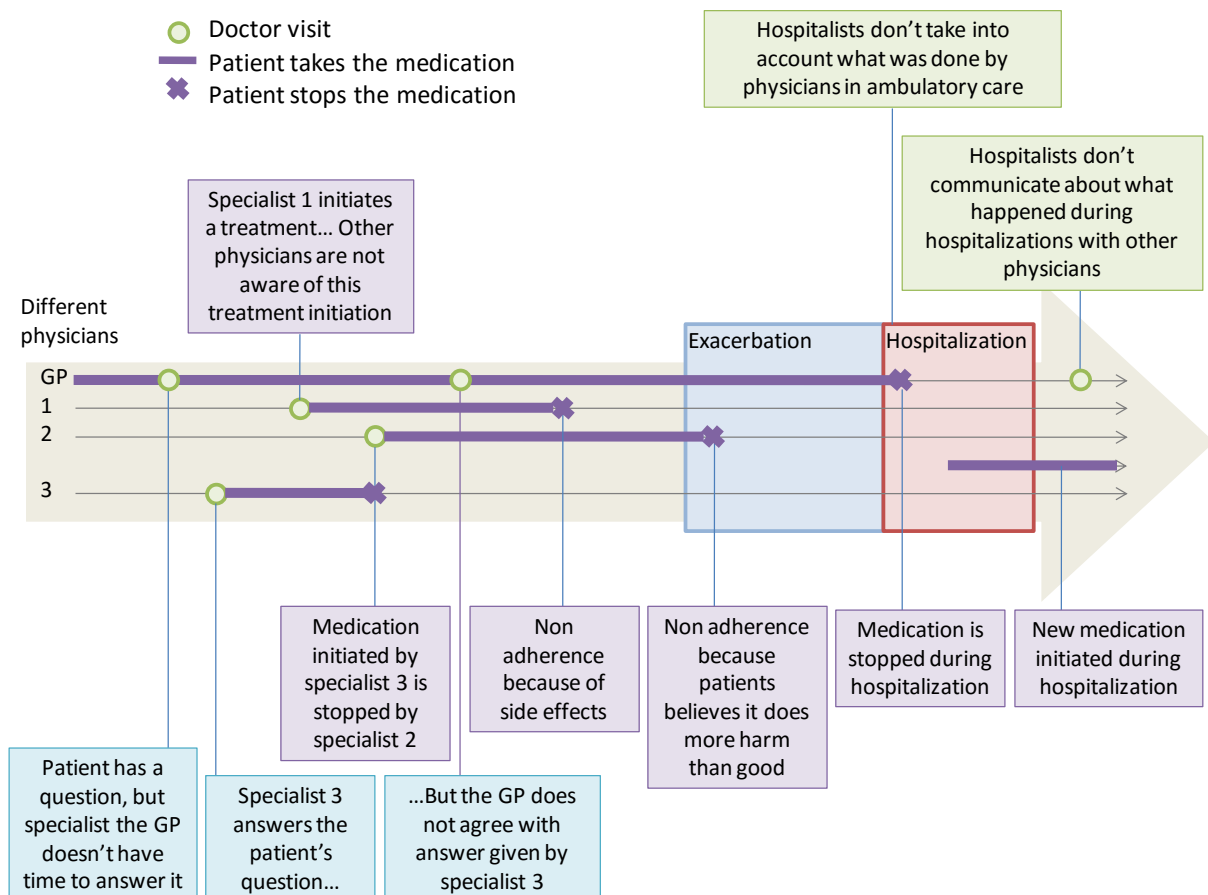


Figure 3 : Comment le manque de coordination entre les médecins peut rendre le suivi complexe. A partir d'un patient fictif caricatural.

Les recommandations de pratique clinique sont définies comme « des propositions développées méthodiquement pour aider le praticien et le patient à rechercher les soins les plus appropriés dans des circonstances cliniques données » [26]. Elles sont développées par la synthèse rigoureuse de l'état de l'art et des données de la science à un temps donné sur un problème donné. Cependant, leur application à des patients multimorbides révèle leurs limites.

La première limite vient de la manière même dont les recommandations sont développées, c'est à dire par l'étude systématique des études, en particulier des essais cliniques, sur une maladie donnée. Or, les patients inclus dans ces essais sont rarement multimorbides [27]. Une revue des essais contrôlés randomisés publiés en 1999-2000 a montré que le rapport entre le nombre de patients inclus et ceux évalués pour éligibilité, pouvait aller jusqu'à 1/68, du fait des comorbidités et traitements concomitants [28]. En cardiologie, on estime que seulement 18% à 25% des patients souffrant d'insuffisance cardiaque correspondent aux critères d'inclusion des essais « clés » guidant leurs traitements [29]. Nous pouvons donc déjà nous interroger sur les preuves à notre disposition pour prendre en charge les patients multimorbides.

Ensuite, les recommandations actuelles ne tiennent pas –ou peu- compte des autres maladies et traitements du patient. Un traitement recommandé pourrait être moins bénéfique, voire délétère, chez des patients avec plusieurs maladies. En étudiant les possibles interactions entre les recommandations de pratique clinique pour le traitement de l'insuffisance cardiaque et de 18 maladies fréquemment associées, Muth et al. a mis en évidence 247 interactions possibles: entre maladies, entre maladies et traitements ou entre plusieurs traitements. Les interactions entre traitements et maladies pouvaient par exemple concerner l'utilisation de bêtamimétiques dans le traitement de l'asthme chez les patients avec une insuffisance cardiaque. Les interactions entre plusieurs traitements pouvaient concerner l'utilisation du millepertuis dans le traitement de la dépression et les traitements de l'insuffisance cardiaque [30]. A l'heure actuelle, les médecins ne disposent que de très peu de données pour décider quels traitements prescrire ou ne pas prescrire en cas d'interactions. Et ceci ne représente que la partie visible de l'iceberg: la plupart des

interactions médicamenteuses, en particulier celles concernant 3 -ou plus- médicaments sont méconnues et imprévisibles.

L'ensemble de ces problèmes est aggravé par la prolifération de ces recommandations. En France, en 2014, on dénombre 95 recommandations de la Haute Autorité de Santé. Ceci sans compter les multiples recommandations de sociétés savantes nationales ou internationales. L'addition de ces recommandations multiples aboutit souvent à des régimes thérapeutiques complexes. En considérant un patient comme la simple addition de ses problèmes de santé, un médecin devrait prescrire 12 médicaments différents (19 comprimés) à 5 moments de la journée, pour une patiente de 79 ans avec de l'ostéoporose, de l'arthrose, du diabète de type 2, de l'hypertension et une bronchite chronique [31] (**Table 1**). Et ceci ne concerne pas uniquement les traitements médicamenteux : les recommandations préconisent l'implication du patient qui va développer des compétences pour prendre en charge sa maladie, la surveiller et changer ses habitudes de vie.

Time	Medications†	Other
7:00 AM	Ipratropium metered dose inhaler 70 mg/wk of alendronate	Check feet Sit upright for 30 min on day when alendronate is taken Check blood sugar
8:00 AM	500 mg of calcium and 200 IU of vitamin D 12.5 mg of hydrochlorothiazide 40 mg of lisinopril 10 mg of glyburide 81 mg of aspirin 850 mg of metformin 250 mg of naproxen 20 mg of omeprazole	Eat breakfast 2.4 g/d of sodium 90 mmol/d of potassium Low intake of dietary saturated fat and cholesterol Adequate intake of magnesium and calcium Medical nutrition therapy for diabetes‡ DASH‡
12:00 PM		Eat lunch 2.4 g/d of sodium 90 mmol/d of potassium Low intake of dietary saturated fat and cholesterol Adequate intake of magnesium and calcium Medical nutrition therapy for diabetes‡ DASH‡
1:00 PM	Ipratropium metered dose inhaler 500 mg of calcium and 200 IU of vitamin D	
7:00 PM	Ipratropium metered dose inhaler 850 mg of metformin 500 mg of calcium and 200 IU of vitamin D 40 mg of lovastatin 250 mg of naproxen	Eat dinner 2.4 g/d of sodium 90 mmol/d of potassium Low intake of dietary saturated fat and cholesterol Adequate intake of magnesium and calcium Medical nutrition therapy for diabetes‡ DASH‡
11:00 PM	Ipratropium metered dose inhaler	
As needed	Albuterol metered dose inhaler	

Table 1 : Traitements recommandés pour une patiente de 79 ans avec de l’ostéoporose, de l’arthrose, du diabète de type 2, de l’hypertension et une bronchite chronique. Adapté de la source [31]

Toutes les interventions recommandées le sont-elles avec « la même force » ? A l’heure actuelle, les traitements recommandés sont hiérarchisés en fonction du niveau de preuve de l’information. Les médecins n’ont pas accès, de manière lisible, à l’impact des traitements en termes de bénéfice ou d’effets indésirables pour le patient [15]. Par exemple, dans les recommandations du traitement de l’hypercholestérolémie de « l’American College of Cardiology (ACC) / American Heart Association (AHA) », publiée en 2013, le traitement avec une statine est recommandé, en prévention

primaire, chez les patients diabétiques âgés de 40 à 75 ans [32]. En fonction des autres facteurs de risque cardiovasculaires, le risque cardiovasculaire fatal à 10 ans pour ces patients peut être <5% [33]. Ce traitement est-il alors véritablement recommandé? Ceci est d'autant plus discutable que respectivement 50% et 75% des patients de plus de 66 ans, traités en prévention primaire par une statine, arrêtent le traitement après 1 et 2 ans.

Comparés aux pathologies aiguës dont le modèle a servi au développement des recommandations pour les maladies chroniques, les manifestations cliniques et biologiques des maladies chroniques, les objectifs de traitement, et les thérapeutiques employées, sont davantage dépendants des individus et de leurs contextes [19]. Les recommandations devraient donc prendre comme critères de traitement, des notions importantes pour le patient : pour des patients jeunes cela pourrait être la survie ; pour des patients plus âgés et fragiles, cela pourrait être de préserver la fonction ou le rôle social et améliorer la qualité de vie [34].

1.3. Fardeau du traitement

La prise en charge des maladies chroniques est par définition de longue durée, principalement à domicile, en dehors de structures médicales. Malgré lui, le patient désirant se soigner devra s'impliquer dans sa santé [21]. Dans une étude ethnographique portant sur des patients diabétiques, aux Pays Bas, Mol et al. a montré que les soignants imaginaient l'implication du patient comme leur participation aux décisions concernant leur santé. Les patients, quant à eux, décrivent cela comme « un processus sans fin, plein de surprises, devant être adapté à chaque instant aux réalités d'une vie désordonnée » [35].

Cette notion n'est pas nouvelle : déjà en 1982, Bury et al. décrivait comment la prise en charge d'une maladie chronique pouvait perturber la « biographie » des individus, modifiant profondément leur vie quotidienne [36]. Dans cette étude qualitative à partir d'entretiens semi-structurés auprès de 30 patients souffrant de polyarthrite rhumatoïde, l'auteur avait décrit comment la maladie : 1) perturbait la vision qu'avaient les patients d'eux-mêmes ; 2) perturbait les moyens qu'avaient les patients pour expliquer ce qui leur arrivait ; et 3) entraînait une mobilisation des ressources des patients pour faire face à ces changements. Concernant ce dernier point, les auteurs insistaient sur les difficultés des patients à intégrer les handicaps liés à la maladie dans leurs vies: planifier les moindres actions ou déplacements pouvaient devenir un défi. Ceci était à l'origine d'un isolement social où les patients se cantonnaient à des « terrains connus et familiers ».

En 1985, Corbin et Strauss ont enrichi ce concept en décrivant pour la première fois, la prise en charge des maladies chroniques au domicile comme un travail pour les

patients [37]. Les auteurs se sont appuyés sur les interviews de 60 patients souffrant de maladies chroniques variées. Cet article a permis de mettre en lumière la nature particulière des maladies chroniques: leur gestion prend place dans la vie quotidienne des patients. Elle repose sur un « équilibre relatif » entre le travail que doivent faire les patients pour se soigner et leurs autres rôles (familial, professionnel, etc.). « *Essayer de satisfaire toutes les demandes de la vie est comme marcher sur une corde raide* ». Dans leur étude, les patients faisaient des efforts pour maintenir un équilibre entre l'énergie, le temps et les finances investies dans chaque aspect de leurs vies. La principale difficulté rencontrée était la quantité de temps nécessaire pour incorporer les demandes imposées par la maladie dans la vie quotidienne. Par exemple, s'ils respectaient l'ensemble des conseils de leurs médecins, les patients diabétiques de type 2 devraient passer en moyenne 122 à 143 minutes par jour à s'occuper de leur maladie et de leurs traitements (**Table 2**) [38].

Estimated time required for recommended care*	
Task	Minutes/day
<i>ADA recommendations</i>	
Home glucose monitoring	3
Record keeping	5
Taking oral medication	4
Foot care	10
Oral hygiene, flossing	1
Problem solving	12
Meal planning	10
Shopping	17
Preparing meals	30
Exercise	30
ADA SUBTOTAL	122
<i>Other desirable self-care</i>	
Monitoring blood pressure	3
Stress management	10
Support group	2
<i>Administrative tasks</i>	
Phoning educators, doctors	1
Scheduling appointments	1
Insurance dealings	2
Obtaining supplies	2
TOTAL TIME	143

Table 2 : Temps nécessaire pour suivre l'ensemble des recommandations du traitement du diabète. Adapté de la source [38]

Dans une étude transversale auprès de 2540 patients australiens de plus de 50 ans, Jowsey et al. a montré que les patients déclaraient consacrer un temps médian allant de 5 à 16 heures à s'occuper de leur santé (en excluant le temps dédié à l'exercice physique). Au maximum, les patients rapportaient consacrer entre 2 et 3h par jour à s'occuper de leur santé [39].

Outre l'investissement de temps, certaines tâches sont peu pratiques, désagréables, ou sont accompagnées d'effets secondaires réels ou craints [40]. Par exemple, il a été montré que les patients font des efforts pour s'adapter aux changements fréquents de

forme et de couleur de leurs comprimés – environ 29% des patients ont au moins 1 changement de forme, taille ou couleur dans leur traitement dans l'année suivant son initiation [41]-.

Si les contraintes en termes de temps et d'inconvénients d'un traitement donné peuvent être semblables entre deux patients avec la même maladie, elles peuvent avoir un impact très différent sur leurs vies [1]. L'impact sur la qualité de vie de tout ce que fait un patient pour se soigner est appelé « fardeau du traitement ». Pour une même tâche, le fardeau peut varier en fonction des capacités mentales, psychologiques, physiques, socio-économiques (i.e. langue, éducation, niveau de vie, présence d'un entourage etc.) et de l'entourage des patients [42].

Des expériences ont montré que les patients sont généralement plus observants pour les recommandations et traitements qui réduisent plus « le fardeau de la maladie » (symptômes, croyance en sa gravité, complications futures, etc.) qu'ils ne créent « un fardeau du traitement » [43]. L'association entre « lourdeur » du traitement (i.e. nombre de prises par jour ou nombre de comprimés), complexité du traitement, effets secondaires perçus, coût financier du traitement et observance thérapeutique est connue depuis longtemps [44]. Pour reprendre l'exemple précédent, les changements de couleur et de forme des comprimés sont associés à l'arrêt du traitement (OR=1.34 IC à 95% [1.12 - 1.59] pour la couleur et OR=1.66 IC à 95% [1.43 - 1.94] pour la forme des comprimés). Ce défaut d'observance pourrait entraîner un mauvais contrôle de la maladie et conduire à une augmentation du nombre d'hospitalisations ou de la mortalité [45, 46]. Cependant, il est probablement plus exact de dire que l'observance est associée au fardeau créé par les tâches imposées plutôt que par la quantité de

tâches elles-mêmes. Malheureusement, les médecins sont souvent ignorants des difficultés que peuvent avoir leurs patients à intégrer leurs traitements dans leurs vies quotidiennes : face à une maladie déséquilibrée ou une aggravation des symptômes, leur principale réponse est souvent d'intensifier le traitement et donc d'aggraver le fardeau du traitement [47].

Afin d'améliorer la prise en charge des patients multimorbides, il est donc nécessaire de repenser la prise en charge des maladies chroniques afin d'adapter leur traitement et leur suivi en fonction de leurs capacités: c'est-à-dire pratiquer une médecine ayant le moins d'impact possible sur leur vie quotidienne, une « Minimally Disruptive Medicine »[1]. Pour arriver à ce but, les médecins et chercheurs ont besoin de moyens pour identifier les patients souffrant du fardeau de leur traitement. Ces méthodes d'évaluation doivent dépasser la simple évaluation des effets secondaires du traitement pour évaluer comment l'observance au long cours de traitements peut perturber la vie des patients [48].

2. Objectif de la thèse

L'objectif de cette thèse était d'élaborer des outils pour mesurer le fardeau du traitement, utilisables chez des patients souffrant de plusieurs maladies chroniques. Ces outils devaient permettre de distinguer, entre plusieurs patients avec la même maladie et le même traitement, ceux qui l'avaient intégré dans leur routine quotidienne de ceux qui luttent pour réaliser toutes les tâches demandées.

Lorsque nous avons débuté cette thèse, aucun outil n'existait pour mesurer le fardeau du traitement de manière globale. Notre premier objectif était donc de développer cet outil. Mesurer un concept subjectif tel que le fardeau du traitement est complexe. Il est nécessaire de s'assurer que l'outil mesure bien le concept qu'il est censé mesurer – c'est-à-dire que l'outil est valide- et que cette mesure est dénuée d'erreur – c'est à dire qu'elle est fiable-. Dans notre premier travail, nous avons élaboré un instrument pour mesurer le fardeau du traitement chez des patients souffrant de plusieurs maladies chroniques et nous avons étudié les caractéristiques de cet outil. Dans notre second travail, nous avons adapté notre outil afin de l'utiliser dans des pays anglophones.

Notre deuxième objectif était de décrire le concept de fardeau du traitement du point de vue du patient: que font les patients pour s'occuper de leur santé ? Quels facteurs aggravent le fardeau du traitement et quelles conséquences ce dernier peut-il avoir sur la vie des patients ? En effet, jusqu'ici, le concept de fardeau du traitement n'avait été décrit que dans des pathologies précises ou pour des traitements donnés. Pour comprendre ce phénomène de manière globale, chez des patients souffrant de

plusieurs maladies chroniques, nous avons réalisé, lors de notre troisième travail, une étude qualitative internationale au cours de laquelle nous avons interrogé des patients sur le fardeau du traitement. Afin de recruter des patients distants géographiquement, nous avons réalisé une étude en ligne. Ceci nous a permis, dans un quatrième travail méthodologique, de travailler sur le concept de saturation des données dans les études qualitatives en ligne.

3. Mesurer le fardeau du traitement

Le premier objectif de cette thèse était de développer des outils pour mesurer le fardeau du traitement chez des patients souffrant de plusieurs maladies chroniques. Dans ce chapitre, nous aborderons tout d'abord les principes des mesures en santé et les critères de qualité des mesures rapportées par les patients. Ensuite, nous présenterons notre premier travail consistant au développement et à la validation du premier instrument pour mesurer le fardeau du traitement chez des patients souffrant de plusieurs maladies chroniques.

3.1. Mesurer la santé et évaluations rapportées par le patient

Les médecins s'appuient sur de nombreuses mesures cliniques pour soigner leurs patients: taille, poids, pression artérielle, valeurs biologiques ou radiologiques, etc. Cependant, pour certains paramètres, les informations les plus fiables viennent directement du point de vue des patients, par exemple : la qualité de vie, la douleur ressentie par le patient, etc. Pour ces paramètres-là, le médecin aura souvent une représentation inexacte des perspectives du patient [49]. Par exemple, lorsque l'on compare l'évaluation de la douleur de patients en postopératoire selon le point de vue du patient ou du médecin, on note que les médecins sous-estiment systématiquement la douleur des patients et surestiment l'efficacité des antalgiques prescrits [50].

Pour évaluer des phénomènes « inobservables », les médecins ont donc besoin de mesures intermédiaires « observables » corrélées au phénomène à mesurer. Par exemple, pour mesurer la température, nous pourrions utiliser une colonne de

mercure dans un tube de verre. Pour des phénomènes subjectifs, nous utilisons des questionnaires remplis par les patients. Ces mesures provenant du patient (en anglais, « Patient Reported Outcome ») doivent, comme le thermomètre, être le plus possible dénuées d'erreur de mesure. L'expérience clinique a montré que le contenu explicite du discours d'un malade ne traduit pas toujours le vécu intime de sa maladie [51]. Pour déterminer à quel point les informations produites par un questionnaire sont une estimation dénuée d'erreur du phénomène à mesurer, une méthode est de considérer ce problème sous l'angle de la théorie classique des tests [52].

3.2. Théorie classique des tests

La théorie classique des tests repose sur le principe que la mesure obtenue en utilisant un instrument (X) est déterminée par l'état « vrai » du phénomène à mesurer (T) + une erreur de mesure (E). Soit : $X = T + E$.

Prenons l'exemple d'un phénomène subjectif tel que la douleur. Si nous demandions à un patient de quantifier sa douleur, entre « 0 » et une valeur maximale arbitraire, ce dernier interpréterait le phénomène ressenti et construirait une réponse en fonction de cette interprétation. La mesure obtenue serait alors sujette à deux types d'erreurs : une erreur aléatoire « incompressible » ; et une erreur relative à la « performance » de la question posée. Pour évaluer la performance de la question posée, il faut se souvenir que, par définition, la mesure obtenue est une mesure intermédiaire d'un phénomène non observable. De fait, elle co-varie avec le phénomène à mesurer. Un index d'association entre ces deux informations permettrait donc d'évaluer l'imprécision associée à l'instrument [52]. Mais, comment obtenir l'information sur le vrai phénomène ou « vrai score », inobservable ?

Pour répondre à ce problème, considérons que nous ne disposons pas d'un unique item, mais de plusieurs items mesurant le même concept. Les relations entre ces items nous donnent alors des informations supplémentaires sur le « vrai score » à mesurer. En effet, si deux items X et Y sont strictement parallèles (c'est-à-dire, si 1) ces 2 items mesurent le même concept, c'est-à-dire sont associés au même « vrai score » ; 2) l'erreur associée à l'item X est indépendante de l'erreur associée à l'item Y et au vrai score ; et 3) les « performances » des items sont similaires) alors la corrélation entre ces deux items est égale au produit des corrélations de chaque item avec le « vrai score » (T). Soit $r_{XY} = (r_{XT})(r_{YT})$. Etant donné que $r_{XT} = r_{YT}$ (condition du parallélisme), alors $r_{XY} = (r_{XT})(r_{YT}) = r_{XT}^2 = r_{YT}^2$.

De fait, à partir d'informations sur les corrélations entre deux items, nous pouvons déterminer les corrélations entre chaque item et le « vrai score », et donc déterminer l'erreur de mesure associée à chaque item. Les conditions énoncées précédemment sont difficiles à vérifier en pratique. Des modèles moins stricts ont été développés et permettent d'élargir ces conclusions à d'avantages de situations [53]. Nous n'entrerons cependant pas dans le détail de ces modèles.

Comme nous venons de le voir, la théorie classique des tests repose sur le fait que de multiples items, mesurant le même concept, vont s'agréger en « échelles » et que la combinaison des mesures de multiples items isolés va alors donner une information plus « fiable », du vrai score que celle provenant de chaque item pris séparément.

Mais la fiabilité d'une mesure ne suffit pas. Il faut également s'assurer que l'ensemble des items d'un instrument est bien associée au phénomène non observable, c'est la « validité ».

3.3. Validité et fiabilité

Nous allons résumer les principales définitions et concepts sous-tendant les termes de validité et de la fiabilité des mesures en santé.

Le « *cadre conceptuel* » d'un instrument décrit l'ensemble des théories, lois, concepts et données scientifiques formant le contexte dans lequel l'instrument et sa mesure prennent place. Ce cadre conceptuel fait le lien entre le phénomène (ou concept) à mesurer et les questions (appelées « items ») de l'instrument.

La « *validité* » d'un instrument est sa capacité à mesurer ce qu'il est censé mesurer. Deux questions forment le concept de validation : « Que mesure l'instrument ? » et « Que vaut la mesure ? » [51]. Le premier objectif de la validation est de produire un ensemble de preuves que : 1) il existe une adéquation entre le cadre conceptuel de l'instrument et les items de l'instrument ; et 2) l'ensemble des composantes du/des concept(s) qu'est censé mesurer l'instrument sont représentées par ses items. C'est la « *validité de contenu* ». Le second objectif de la validation est de démontrer que la mesure de l'instrument prend des valeurs cohérentes avec ce que l'on sait du phénomène à mesurer. Pour cela on va comparer la mesure à tester avec des mesures existantes, déjà validées. C'est la « *validité convergente* » (appelée aussi *validité concourante* ou *critérielle*). Si aucune autre mesure de ce concept n'existe, il faut alors chercher à vérifier des hypothèses définies à priori sur le concept à mesurer, par exemple en prédisant les réponses de groupes de participants différents sur des caractéristiques connues (méthode des « groupe connus »). C'est la « *validité de construit* » [54].

La « *dimensionnalité* » d'un instrument représente le fait qu'il mesure un concept unique ou des concepts multiples (pouvant être décomposés en plusieurs dimensions ou sous-mesures). Par exemple, la mesure du quotient intellectuel en utilisant la « Wechsler Adult Intelligence Scale » [55] est une mesure multidimensionnelle de la compréhension verbale, du raisonnement perceptif, de la mémoire de travail et de la vitesse de traitement. Comme nous l'avons vu précédemment, il est important de s'assurer de la dimensionnalité d'un instrument afin de regrouper ses items en sous-mesures unidimensionnelles, car sinon, il est impossible d'évaluer sa fiabilité.

La « *fiabilité d'un instrument* » représente, comme nous l'avons vu précédemment, la capacité d'un instrument à donner une mesure dénuée d'erreur. De manière générale, les auteurs décomposent la fiabilité entre la « *consistance interne* » et la « *stabilité de la mesure* ». La consistance interne représente la capacité des différents items d'un instrument mesurant le même concept, à donner un même score. En effet, si deux items (X et Y) sont corrélés, cela signifie qu'ils ont quelque chose en commun responsable de cette corrélation. Ainsi, une partie du score de l'item X est déterminée par le score de l'item Y. Cependant, à moins que la corrélation entre les deux items soit parfaite, chaque item a également une part de variabilité qui lui est propre. Cette considération peut être généralisée pour un ensemble d'items mesurant un même concept. La variabilité de l'ensemble de ces items pourra donc être décomposée en une part commune et une part dépendante de chaque item. Plus grande sera la part commune et plus les items mesureront le même concept et seront un reflet du « vrai score ». La consistance interne d'une mesure est une estimation de la part de variabilité commune que partagent les items entre eux. La stabilité de la mesure, quant

à elle, représente la capacité de l'instrument à donner des résultats similaires lorsqu'il est administré à des moments différents, si le concept à mesurer n'a pas changé.

La « *sensibilité au changement* » d'un instrument est définie comme sa capacité à détecter des changements « cliniquement importants », même si ces changements sont faibles [56]. Un instrument idéal devrait donner des mesures similaires pour les patients chez qui l'objet de la mesure ne change pas (reproductibilité) mais devrait aussi pouvoir détecter des modifications cliniques, même faibles (sensibilité au changement).

« *L'interprétabilité d'un instrument* » représente sa capacité à permettre à ses utilisateurs de mettre un sens clinique sur deux mesures différentes. Sont-elles uniquement le fait d'erreurs de mesure ou s'agit-il d'une différence clinique de l'objet de mesure ? Ce concept est généralement lié à celui de « *différence clinique minimale* ». Cette différence correspond à la plus petite différence de mesure ayant un sens pour les patients ou leurs proches et qui mènerait le patient ou le clinicien à considérer un changement dans la prise en charge.

« *Le fardeau de l'enquêteur/répondant* » représente le temps et l'effort demandés aux participants ou aux enquêteurs pour répondre ou administrer l'instrument.

3.4. Importance de l'évaluation de la validité et de la fiabilité

Les mesures provenant du patient sont devenues des critères de jugement aussi importants que les mesures issues d'examen cliniques ou complémentaires dans les essais cliniques. Par exemple, une analyse transversale du registre d'essai Clinicaltrials.gov en 2007, a montré que parmi 17,704 essais interventionnels de phase 3

enregistrés, 2,481 (14%) utilisaient au moins une mesure subjective comme critère de jugement [57]. De plus en plus, les chercheurs tentent de replacer le patient au centre de l'évaluation de sa santé et utilisent des mesures provenant du patient comme critères d'efficacité et de qualité des traitements. Aux Etats-Unis, ceci a abouti à la création du « Patient-Centered Outcomes Research Institute » (PCORI) dans le « US Patient Protection and Affordable Care Act » en 2010 [58].

L'utilisation d'une mesure mal validée ou peu fiable influe directement sur la mesure et les résultats des études, et donc, sur la prise en charge des patients. A partir d'un échantillon de 300 essais contrôlés randomisés chez des patients schizophrènes, Marshall et al. a constaté que l'utilisation d'échelles non validées était significativement associée à un résultat positif dans l'étude (RR=1.37 IC à 95% [1.12-1.68]) [59]. Etant donné qu'il n'y avait aucune raison de penser que ces échelles sont meilleures pour mesurer l'effet du traitement, nous pouvons supposer qu'elles rapportaient des résultats erronés du « vrai score ». Bien qu'il puisse y avoir plusieurs explications à ce résultat, il n'en demeure pas moins essentiel d'utiliser des méthodes adéquates afin de s'assurer de la validité et la fiabilité des instruments de mesure en santé, y compris les instruments de mesure subjective.

3.5. Critères de qualité des instruments de mesures provenant du patient

De nombreuses recommandations ont été développées afin d'assurer la qualité des instruments de mesure en santé. Nous en présenterons deux exemples : 1) les recommandations éditées par la « Food and Drug Administration » (FDA) [60] qui énoncent les critères de qualité utilisés par l'agence lors de l'examen des demandes

d'autorisations de commercialisation des médicaments sur le territoire Américain; et 2) les recommandations « minimales » de « l'International Society for Quality of Life Research » (ISOQOL) [61], conçues par la synthèse d'experts des 28 principales recommandations pour le développement ou l'évaluation de mesures provenant du patient. Ces deux recommandations reposent sur les mêmes points : 1) présence d'un cadre conceptuel décrit de manière précise ; 2) évaluation de la validité de contenu et de construit de l'instrument; 3) évaluation de la fiabilité de l'instrument ; 4) évaluation de l'interprétabilité des scores ; 5) évaluation des éventuelles traductions ; et 6) évaluation du fardeau du répondant (**Table 3**).

Un instrument de bonne qualité doit comprendre une définition précise du ou des phénomène(s) à mesurer et des conditions d'utilisation de l'instrument (i.e. pour quelle population, pour quelles pathologies, etc.). Ce cadre conceptuel est généralement défini par l'étude de la littérature et d'interviews avec des experts du sujet et des patients. C'est à ce stade que sont également décidées : 1) la longueur du questionnaire (avec une idée approximative du nombre d'items), 2) les méthodes d'administration (i.e. hétéro-questionnaire ou auto-questionnaire, contexte d'utilisation, etc.), 3) les options de réponse, et 4) la méthode de calcul de score.

La qualité de la validité de contenu d'un instrument est déterminée par l'adéquation entre le cadre conceptuel et les différents items de l'instrument. Les deux recommandations prônent l'élaboration d'une liste d'items en deux étapes : 1) recherche d'une liste exhaustive d'items à partir de la littérature, d'interviews avec des médecins et patients, et de questionnaires déjà existants; et 2) réduction du nombre d'items en ne conservant que les plus importants et les plus performants (e.g.

répartition appropriée des réponses, faible taux de réponses manquantes ou incohérentes, etc.). Les standards minimums de l'ISOQOL insistent sur l'importance de décrire : 1) les méthodes et sources employées pour établir la liste initiale d'items; 2) les caractéristiques des patients ayant évalué la clarté et compréhensibilité des items (et notamment leurs différences avec les patients à qui l'instrument est destiné); et 3) une justification de la période de rappel de l'instrument.

La qualité de la validité de construit est déterminée par l'étude des relations entre la mesure issue des items de l'instrument et d'autres mesures. Pour cela, le chercheur va poser des hypothèses sur ces relations, avant toute collecte de données et va ensuite les vérifier expérimentalement. L'instrument est jugé valide si les relations postulées a priori sont exactes. Terwee et al. recommande de s'assurer que 75% des hypothèses présumées sont vérifiées dans des sous-groupes d'au moins 50 participants avant de déclarer que la validité de construit a été évaluée [62].

L'évaluation de la fiabilité de l'instrument est le seul point sur lequel les deux recommandations présentées sont divergentes. La FDA considère qu'un instrument n'est fiable que lorsqu'à la fois: l'évaluation de sa consistance interne et la stabilité de sa mesure ont été évaluées. De son côté, l'ISOQOL considère comme standard minimum l'évaluation de la seule consistance interne. Pour évaluer la stabilité de la mesure, les deux recommandations préconisent une méthode de « test-retest ». Dans cette méthode, l'instrument est administré à deux moments différents: si le concept à mesurer n'a pas changé, alors les deux mesures devraient être concordantes. Le choix de l'intervalle de temps doit être ni trop long (afin d'éviter que le concept ne change), ni trop court (afin d'éviter un biais de mémorisation). La mesure de choix afin

d'évaluer la concordance entre les mesures est le coefficient de corrélation intra-classe (ICC) [62, 63]. Pour évaluer la consistance interne d'une échelle, les recommandations préconisent le coefficient alpha de Cronbach. Un coefficient alpha faible indique une faible corrélation entre les items d'une échelle qui doit faire discuter l'idée de les combiner en un score unique. A l'inverse, un coefficient alpha élevé évoque de fortes corrélations entre items. Les auteurs jugent la consistance interne de l'échelle acceptable lorsque alpha est compris entre 0.7 et 0.9 [62, 64].

Dans les recommandations, la sensibilité au changement est considérée comme acceptable lorsque l'instrument apporte la preuve qu'il est sensible à l'élévation comme à la diminution du concept à mesurer, et ce, pour l'ensemble des valeurs possibles. Une mesure possible est de calculer l'aire sous la courbe (AUC) de performance du test, ou « Receiver Operating Characteristic » (ROC) après avoir défini si les patients avaient changé d'état à l'aide d'une mesure externe. Une aire sous la courbe >0.7 est considérée comme adéquate [62]. L'évaluation de la sensibilité au changement ne fait pas partie des standards minimums de l'ISOQOL [61].

La présence d'une documentation permettant d'interpréter les scores était considérée comme un standard minimum d'un instrument de mesure par l'ISOQOL.

Pour conclure, les mesures provenant du patient occupent une place importante dans l'évaluation de la santé, aussi bien en clinique qu'en recherche. Leur développement et leur évaluation nécessitent une approche rigoureuse.

Measurement Property	Type	What Is Assessed?	FDA Review Considerations
Reliability	Test-retest or intra-interviewer reliability (for interviewer-administered PROs only)	Stability of scores over time when no change is expected in the concept of interest	<ul style="list-style-type: none"> • Intraclass correlation coefficient • Time period of assessment
	Internal consistency	<ul style="list-style-type: none"> • Extent to which items comprising a scale measure the same concept • Intercorrelation of items that contribute to a score • Internal consistency 	<ul style="list-style-type: none"> • Cronbach's alpha for summary scores • Item-total correlations
	Inter-interviewer reliability (for interviewer-administered PROs only)	Agreement among responses when the PRO is administered by two or more different interviewers	<ul style="list-style-type: none"> • Interclass correlation coefficient
Validity	Content validity	Evidence that the instrument measures the concept of interest including evidence from qualitative studies that the items and domains of an instrument are appropriate and comprehensive relative to its intended measurement concept, population, and use. Testing other measurement properties will not replace or rectify problems with content validity.	<ul style="list-style-type: none"> • Derivation of all items • Qualitative interview schedule • Interview or focus group transcripts • Items derived from the transcripts • Composition of patients used to develop content • Cognitive interview transcripts to evaluate patient understanding
	Construct validity	Evidence that relationships among items, domains, and concepts conform to <i>a priori</i> hypotheses concerning logical relationships that should exist with measures of related concepts or scores produced in similar or diverse patient groups	<ul style="list-style-type: none"> • Strength of correlation testing <i>a priori</i> hypotheses (discriminant and convergent validity) • Degree to which the PRO instrument can distinguish among groups hypothesized <i>a priori</i> to be different (known groups validity)
Ability to detect change		Evidence that a PRO instrument can identify differences in scores over time in individuals or groups (similar to those in the clinical trials) who have changed with respect to the measurement concept	<ul style="list-style-type: none"> • Within person change over time • Effect size statistic

Table 4 : Résumé des critères de qualité de la FDA pour évaluer les instruments utilisés dans les essais cliniques.

3.6. Mesures existantes de la satisfaction du patient vis-à-vis de son traitement

Il n'existait pas, lorsque nous avons débuté ce travail, de mesure du fardeau du traitement utilisable quel que soit le contexte de maladie ou de traitement du patient. De nombreux instruments (ou sous-mesures d'instruments) mesuraient des concepts proches tels que: 1) la satisfaction du patient vis-à-vis de son traitement ou 2) « l'impact du traitement » sur le patient dans des pathologies spécifiques ou pour des traitements particuliers.

Une revue systématique des instruments évaluant ces concepts dans 3 maladies chroniques (diabète, insuffisance rénale et insuffisance cardiaque) a retrouvé 57 instruments [65]. Ces instruments évaluaient les domaines suivant: 1) impact émotionnel du traitement; 2) conflit familial; 3) commodité du traitement; 4) auto-surveillance; 5) impact social et professionnel; 6) planification; 7) effets secondaires ; 8) problèmes associés au régime; 9) « fardeau du traitement »; 10) dérangement associé à l'utilisation d'équipements et 11) aspects financiers. Cette revue systématique n'a identifié aucun instrument évaluant ces aspects de manière générale.

La qualité méthodologique du développement des instruments retrouvés était hétérogène. Sur le plan conceptuel, 38 (67%) avaient été développés en impliquant des patients. Concernant la fiabilité de ces instruments, 46 (81%) comprenaient une évaluation de la consistance interne mais seulement 9 (15%) avaient évalué la stabilité de la mesure par une méthode de test-retest. Concernant leur validité, 32 (56%) comprenaient une étude de la dimensionnalité; 23 (40%) avaient apporté une preuve de validité convergente et 47 (82%) avaient respectivement apporté une preuve de leur

validité de construit par une méthode de « groupes connus ». Enfin, 11 (19%) des instruments apportaient une information sur leur sensibilité au changement. Le détail des instruments et de leurs caractéristiques figure dans la **Table 4**.

Devant l'absence d'instrument et la qualité moyenne des alternatives existantes, nous avons développé et validé un nouvel instrument, pour mesurer le fardeau du traitement.

Instrument	Relevant subscales, items (n)	Patient input	Reliability	Scale analyses	Convergent/discriminant validity	Known-group/concurrent validity	Sensitive to change
Diabetes measures							
Barriers to Adherence Questionnaire ^{56,92}	Diet (4), insulin injection (3), exercise (3), glucose testing (5), total (15)	√	X	NR	NR	√	NR
Barriers to Diabetes Adherence ²⁶	Stress/burnout (4), time pressure/planning (5), stigma (6)	√	√	√	NR	√	NR
Barriers in Diabetes Questionnaire ⁴⁹	Self-control and advice from providers (9), injecting, monitoring and overall self-regulation (10), self-regulation in specific situations (9), total (28)	X	√	√	NR	√	NR
Blood Glucose Monitoring System Rating Questionnaire ^{58,93}	Convenience (NR), interference (NR), blood glucose burden (NR)	X	√	NR	NR	√	X
Continuous Glucose Monitoring Satisfaction Scale ^{61,62,78}	Hassles (20)	X	√	√	NR	√	NR
Diabetes-39 ^{67,94}	Diabetes control (12)	√	√	√	√	√	NR
Diabetes Distress Scale ^{28,73,84,95}	Emotional burden (5), regimen-related distress (5)	√	√	√	√	√	√
DDS-2, DDS-3, and DDS-4 ²⁷	Totals for scales (2-items, 3-items, and 4-items, respectively)	√	√	NR	√	√	NR
Diabetes Family Adherence Measure ³⁹	Coercion (7)	X	√	√	√	√	NR
Diabetes Family Behavior Checklist ^{40,96}	Nonsupportive behaviors (7)	√	X	NR	√	√	√
Diabetes Family Conflict Scale-Revised ^{41,86,97-100}	Direct management (9), indirect management (10), total (19)	√	√	√	√	√	√
Diabetes Family Support and Conflict Scale ⁴²	Family conflict (4)	√	√	√	NR	√	NR
Diabetes Fear of Injecting and Self-testing Questionnaire ^{30,79,91}	Self-injecting fear (15), self-testing fear (15), total (30)*	X	√	√	√	√	NR
Diabetes Health Profile ^{37,82,101,102}	Barriers to activity (NR)	√	√	√	√	√	√
Diabetes Medication Satisfaction Measure ^{46,81,90,103}	Burden (11), symptoms (5)	√	√	√	√	√	√
Diabetes Medication Treatment Satisfaction Tool ⁴⁷	Lifestyle (5), convenience (3), well-being (3)	√	√	√	NR	√	NR
Diabetes-specific Quality of Life Scale ^{48,75,80}	Diet restrictions (9), daily hassles (6)	√	√	√	√	√	√
Diabetes Responsibility and Conflict Scale ⁴³	Conflict (NR)	X	NR	NR	NR	X	NR
Diabetes Self-Care Barriers Assessment for Older Adults ^{60,104}	Glucose monitoring barriers (4), diet regimen barriers (4), exercise barriers (4), total (12)	√	√	√	NR	√	NR
Diabetic Foot Ulcer Scale ^{59,105}	Ulcer care (4), financial (2)	√	√	√	√	√	X
DISABKIDS Diabetes module ^{24,106}	Impact/acceptance (6), treatment (4)	√	√	√	NR	√	NR
General Barriers to Diabetes Self-Management ^{57,107}	Diet (7), exercise (7), glucose testing (7), medication taking (7), general barriers (4), total (31)	√	√	NR	NR	√	NR
Dietary Barriers ^{57,107}	At home (7), food purchase (8), away from home (5), total (27)	√	√	NR	NR	√	NR
Glucose Monitoring Survey ⁶²	Glucose control (7), social complications (6), total (22) (note, 9 items do not load on either subscale)	X	√	√	√	√	NR

Insulin Delivery System Rating Questionnaire ^{50,58,93}	Convenience (NR), interference (NR), blood glucose burden (NR) ⁵	√	√	NR	NR	√	√
Insulin Pen Questionnaire ⁵¹	Convenience (7), facilitation of self-care (6), cost (1)	√	√	NR	NR	√	NR
Insulin Treatment Appraisal Scale ³¹	Negative appraisal (16)	√	√	√	√	√	NR
Insulin Treatment Questionnaire ⁴⁵	Insulin therapy perception (12)	X	√	√	X	X	√
Insulin Treatment Satisfaction Questionnaire ^{52,108-111}	Inconvenience of regimen (5) lifestyle flexibility (3), delivery device satisfaction (6) ⁵	√	√	√	√	√	X
Measure of Invasiveness as a Reason for Skipping Self-Monitoring of Glucose ⁶³	Total (7)	X	√	NR	√	√	NR
Multidimensional Diabetes Questionnaire ⁴⁴	Misguided support behaviors (4)	√	√	√	NR	√	NR
Multidimensional Diabetes Self-Management Checklist ⁵³	Burden of diet self-management (3), burden of injecting insulin (3), burden of glucose monitoring (3), burden of adjusting insulin dose (3)	X	√	NR	NR	√	NR
Patient Satisfaction with Insulin Therapy ^{54,74}	Convenience/ease (10), social comfort (5), global satisfaction (15)	√	√	√	NR	√	√
Pediatric Quality of Life Inventory 3.0 Diabetes module ^{25,112,113}	Treatment barriers (4), treatment adherence (7)	√	X	NR	√	√	NR
Perceived Burden of Diabetes Treatment ¹⁸	Single item on perceived burden of treatment	X	NR	NA	NR	√	NR
Perceived Difficulties in Diabetes Self-Care ²²	Individual items assessing difficulties in insulin treatment (2), glucose monitoring (1), diet (3), exercise (1), smoking (1), daily self-care (1), and self-care during certain occasions (4)	√	NR	NR	NR	√	NR
Perceptions About Medications for Diabetes ²⁹	Schedule flexibility (3), portability convenience (2), regimen inconvenience (5), difficulty remembering medications (2), gastrointestinal side effects (3), weight/edema side effects (3), emotional (10)	√	√	√	NR	√	NR
Perceptions of Insulin Shots and Fingersticks ²³	Injection pain (1), injection fear (1), fingerstick pain (1), fingerstick fear (1)	X	NR	NR	NR	√	X
Personal Diabetes Questionnaire ³⁸	Diet barriers (7), medication barriers (8), monitoring barriers (8), exercise barriers (7)	√	√	NR	NR	√	NR
Practicality/comfort of treatment ¹⁹	Single item on practicality and comfort of treatment	X	NR	NA	NR	NR	NR
Problem Areas in Diabetes ^{32,77,83,114}	Treatment-related (3) and food-related problems (3)	√	√	√	√	√	√
Survey of Treatment Burdens in Diabetes ²²¹	Individual items assessing burden of: oral agents (1), insulin (4), oral + insulin therapy (1), self-monitoring of blood glucose (3), diet (2), pain (1), and interference with activities (1)	X	NR	NR	NR	√	NR
Treatment-Related Impact Measure – Diabetes ⁶	Treatment burden (6), daily life (5), psychological health (8)	√	√	√	√	√	NR
TRIM-Diabetes Device ⁶	Device function (5), bother (3), total (8)	√	√	√	√	√	NR
Treatment Satisfaction Measure for People with Insulin-dependent Diabetes ⁶⁴	Perceived compatibility with lifestyle (PCL) (3)	√	NR	√	NR	√	NR

Instrument	Relevant subscales, items (n)	Patient input	Reliability	Scale analyses	Convergent/discriminant validity	Known-group/concurrent validity	Sensitive to change
Kidney disease measures							
Health Beliefs About Fluid Adherence ³³	Barriers to fluid adherence (7)	X	√	NR	NR	√	NR
Hemodialysis Stressor Scale ^{34,35,85,115}	Physiological, psychosocial, total ^d	X	√	√	NR	X	√
Continuous Ambulatory Peritoneal Dialysis Stressor Scale ³⁵		X	√	NR	NR	NR	NR
Renal Adherence Attitudes Questionnaire ⁶⁵	Attitude towards social restrictions (8), acceptance/lifestyle impact (11)	√	√	√	NR	NR	NR
Satisfaction with Care Questionnaire ⁶⁸	Financial/transportation (7)	√	√	NR	√	NR	NR
Treatment Effects Questionnaire ^{6,36,76}	Total (20)	X	√	NR	√	√	NR
Heart failure measures							
Beliefs about Medicine Compliance Scale ^{55,116}	Medication barriers (6)	√	√	√	NR	NR	NR
Beliefs about Diet Compliance Scale ^{55,116,117}	Diet barriers (5)	√	√	√	NR	X	NR
Dietary Sodium Restriction Questionnaire ^{66,118}	Perceived behavioral control (7) ^a	√	√	√	NR	NR	NR
Perceived Difficulty Affording Health Care ^{20,72}	Single item on economic burden of medical costs	X	NR	NA	NR	√	NR

Notes: ^aMollema et al²⁹ added 4 items to fear of injecting and 4 items to fear of testing subscales; ^bmodified from original scoring²⁹ which scored two subscales, ie, treatment satisfaction (15 items) and treatment interference with activities (11 items); ^cBode et al¹⁰⁹ used a modified scoring procedure for all subscales; ^dnumber of items in the subscales and the overall scale fluctuate across studies. A third subscale ("dependency/restriction") was identified by Murphy et al;¹¹³ ^eoriginal scoring¹¹⁴ recorded only response frequencies of individual items.

Abbreviations: √, satisfactory; X, unsatisfactory or incomplete; NR, not reported; NA, not applicable.

Table 5 : Instruments évaluant la « satisfaction du patient vis-à-vis de son traitement » ou « l'impact du traitement » dans 3 maladies chroniques (diabète, insuffisance cardiaque, insuffisance rénale). Adapté de la source [65].

3.7. Développement d'un instrument de mesure du fardeau du traitement

L'objectif de ce premier travail était de développer et valider un instrument de mesure du fardeau du traitement, « le Treatment Burden Questionnaire » (TBQ), dans une population de patients souffrant de plusieurs maladies chroniques, utilisable à la fois en pratique clinique quotidienne et en recherche clinique.

3.7.1. Méthodes

Nous avons utilisé une méthode classique pour le développement d'instruments mesurant un concept subjectif en santé [62, 66] : 1) élaboration d'un questionnaire à partir d'une revue de la littérature et d'interviews qualitatives de patients; 2) validation de l'instrument obtenu dans une population de patients souffrant de plusieurs maladies chroniques et 3) évaluation de sa fiabilité.

3.7.1.1. Elaboration de l'instrument

Dans un premier temps, nous avons d'abord effectué une revue de la littérature sur MEDLINE en utilisant PubMed afin d'identifier les mesures du fardeau du traitement préalablement existantes. Nous avons trouvé des instruments utilisables chez des patients avec des pathologies ou des traitements spécifiques, mais aucun instrument ne permettait de mesurer le fardeau du traitement de manière globale, chez des patients souffrant de plusieurs maladies chroniques. Cette recherche nous a permis de définir une série de thèmes autour du fardeau du traitement : prises médicamenteuses, auto-surveillance, examens complémentaires, visites chez les médecins, contraintes administratives, difficultés à suivre les conseils pour suivre un régime ou pratiquer des

activités sportives et impact de leur prise en charge médicale sur leurs relations avec les autres.

Dans un second temps, nous avons réalisé 22 interviews qualitatives, à l'aide d'entretiens semi-dirigés, auprès de patients souffrant de maladies chroniques, afin de confirmer et compléter les thèmes identifiés lors de la recherche de la littérature. Ces interviews ont permis d'identifier un nouveau thème du fardeau du traitement : « Le fait de me soigner régulièrement me rappelle que je suis malade ».

Au total, ces deux étapes ont permis l'élaboration d'un questionnaire préliminaire composé de 14 items, évalués par des échelles type likert, graduées de 0 (Aucun impact) à 10 (Impact considérable). Pour chaque question, les patients pouvaient cocher une case « Je ne suis pas concerné ».

3.7.1.2. Evaluation des caractéristiques de l'instrument

L'instrument obtenu a été validé auprès d'une population de patients souffrant d'au moins une maladie chronique, recruté dans des cabinets de médecine générale ou hospitalisés dans des centres hospitaliers universitaires (CHU) en Ile-de-France. La validation s'est déroulée en 4 étapes.

Dans un premier temps, nous avons réduit le nombre de questions, en étudiant les caractéristiques des réponses (réponses manquantes, effet plancher/plafond, corrélation inter-items, etc.) afin d'éliminer des questions redondantes ou peu informatives.

Dans un second temps, nous avons étudié la validité factorielle de notre instrument. Autrement dit, nous avons évalué si notre instrument mesurait un seul concept ou si la mesure conduisait à l'agrégation hétérogène d'items constituant autant de sous-mesures.

Dans un troisième temps, nous avons étudié la validité de construit de notre instrument. Pour cela, nous avons vérifié différentes hypothèses sur le fardeau du traitement. Tout d'abord, nous avons cherché une corrélation négative entre la mesure du fardeau du traitement et la satisfaction du patient vis-à-vis de son traitement, évaluée par le « Treatment Satisfaction Questionnaire for Medication » (TSQM) [67, 68]. Ensuite, nous avons cherché une corrélation positive entre la mesure du fardeau du traitement et 1) le nombre de prises médicamenteuses (nombre de comprimés, par jour, nombre de prises, etc.) ; 2) l'importance du suivi médical (nombre de consultations/an, nombre de spécialistes différents consultés, etc.) ; et 3) le temps passé à s'occuper de sa maladie ou de son traitement. Ces corrélations ont été évaluées à l'aide du coefficient de corrélation de Spearman (r_s). Enfin, nous avons supposé une concordance entre la mesure du fardeau du traitement obtenue en interrogeant le patient, son médecin et un proche.

Dans un dernier temps, nous avons évalué la fiabilité de la mesure à l'aide 1) d'une méthode de test-retest et 2) d'une évaluation de sa consistance interne.

Pour décrire la concordance entre la mesure du fardeau du traitement obtenue en interrogeant le patient, son médecin et un proche et la reproductibilité par le test-retest, nous avons utilisé 1) le coefficient de corrélation intraclass (ICC); et 2) des graphiques de Bland et Altman [69] qui présentent la différence entre deux mesures en fonction de la moyenne de ces deux mesures. Les analyses statistiques ont été effectuées en utilisant SAS v9.2 (SAS Institute, Cary, NC, USA) et R v2.13.1 (<http://www.r-project.org>).

3.7.2. Résultats

Au total, 502 patients (âge moyen 59.3 (\pm 17) ans ; 266 femmes (53.1%)) ont été inclus entre avril 2011 et septembre 2011 lors de la phase de validation du questionnaire ; 257 patients étaient hospitalisés (51.2%). Les patients déclaraient comme principale maladie chronique des pathologies variées telles que le diabète (16.5%), les cancers (6.9%) ou des maladies psychiatriques compensées (1.6%).

Lors de la réduction du nombre d'items, celui portant sur « Les précautions que vous devez prendre pour conserver vos médicaments » a été supprimé car un grand nombre de patients n'étaient pas concernés. Par ailleurs, cet item présentait un effet plancher important (64%).

La validité factorielle de notre instrument a été établie en étudiant le diagramme des valeurs propres d'une analyse factorielle. Nous avons retrouvé une structure unidimensionnelle du questionnaire ; les items de notre instrument, malgré leur diversité, étaient donc le reflet d'un seul et même attribut.

La validation de construit de notre instrument a établi 1) une corrélation négative entre notre instrument et les scores du TSQM (r_s allant de -0.26 à -0.53 ; $p < 0.0001$) et 2) des corrélations positives entre la mesure du fardeau du traitement et les variables évaluant l'importance du traitement (r_s allant de 0.24 à 0.4 ; $p < 0.0001$). Lorsque nous avons étudié la concordance entre la mesure du fardeau du traitement évaluée par le patient, son médecin et un proche, nous avons observé : 1) une concordance moyenne entre la mesure du fardeau du traitement par le patient et par un proche (ICC=0.60 [0.28-0.79], $n=39$); et 2) une mauvaise concordance entre la mesure du fardeau du traitement par le patient et par son médecin (ICC=0.38 [0.29-0.47], $n=396$).

La reproductibilité de la mesure évaluée à l'aide d'une méthode de test-retest était bonne (ICC=0.76 [0.67-0.83], n=211).

3.7.3. Discussion

Au cours de ce travail, nous avons développé le premier instrument de mesure du fardeau du traitement, le « Treatment Burden Questionnaire » (TBQ), dans une population de patients souffrant de plusieurs maladies chroniques. Cet instrument était utilisable quels que soient les maladies ou les traitements du patient.

Les items de cet instrument ont été développés en fonction de l'étude de la littérature, d'avis de cliniciens prenant en charge des patients souffrant de maladies chroniques et d'avis de patients. Il est important de noter la bonne validité de contenu de l'instrument développé puisque les items choisis seront concordants avec : 1) les études qualitatives portant sur le fardeau du traitement réalisées après notre travail [47, 70, 71] et 2) les domaines retrouvés lors de l'étude systématique des instruments évaluant le fardeau du traitement dans 3 maladies chroniques (diabète, insuffisance rénale et insuffisance cardiaque). Par ailleurs, notre étude a permis de montrer que l'instrument développé avait une validité de construit et une fiabilité acceptable, selon les standards internationaux pour le développement des mesures provenant du patient.

Notre étude a également permis de montrer que, les médecins ignoraient souvent les contraintes ressenties par les patients pour suivre l'ensemble de leurs prescriptions: la concordance entre l'évaluation du fardeau du traitement par le patient et celle de son médecin était faible, y compris pour des points spécifiques tels que l'auto-surveillance ou l'observance d'un régime. Ce résultat n'était pas étonnant car, d'une part, le fardeau du traitement est rarement abordé par les patients lors des consultations [72] et,

d'autre part, les médecins se focalisent souvent d'avantage sur des problèmes biomédicaux en occultant le contexte de soin des patients (transport, situation financière, besoin d'un aidant...) [73].

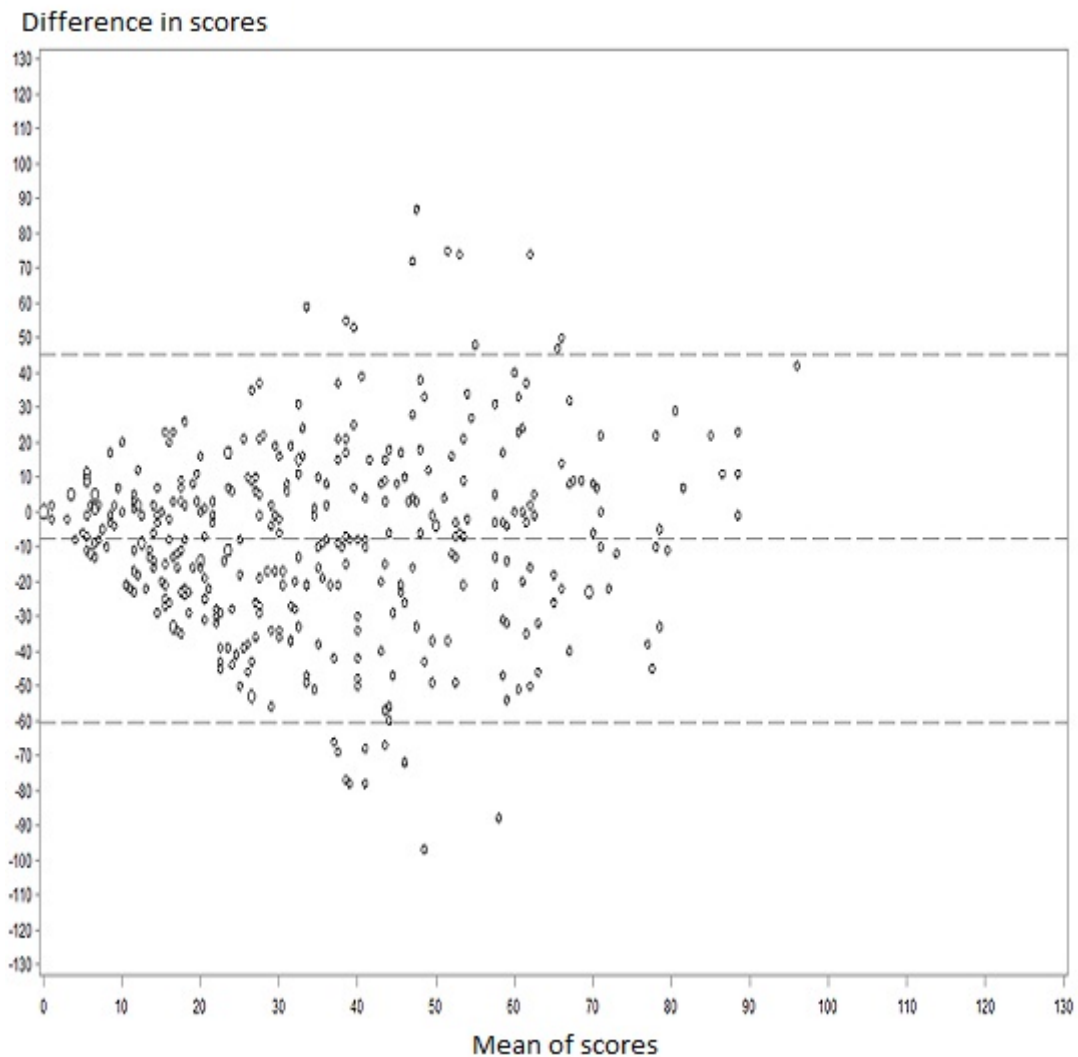


Figure 4 : Graphique de Bland et Altman montrant la concordance entre la mesure du fardeau du traitement évaluée par le médecin et celle évaluée par le patient (n=396 patients). Le graphique montre la différence entre les deux mesures en fonction de la moyenne des deux mesures. Une différence négative signifie que le médecin a surestimé le fardeau de son patient et une différence positive que le médecin a sous-estimé le fardeau de son patient. Les lignes horizontales représentent la différence moyenne entre les deux mesures et les limites d'agrément à 95%. La taille des points est proportionnelle au nombre d'observations.

Les forces de cette étude reposent sur la validation de l'instrument dans un échantillon de patients hospitalisés et ambulatoires, souffrant de pathologies variées et recevant des traitements différents, en France. En ce qui concerne les limites de notre instrument, l'ensemble des items avaient un effet plancher et de nombreux patients avaient répondu « Je ne suis pas concerné » à une ou plusieurs questions. Ce résultat était partiellement attendu car le fardeau du traitement dépend de l'intégration par le patient de sa prise en charge médicale dans sa vie quotidienne ; certains aspects du traitement, devenus une routine pour certains patients (par exemple les prises médicamenteuses quotidiennes), pouvaient ne pas représenter une contrainte pour eux. Ainsi, deux patients avec une même maladie et un même traitement pouvaient ressentir un fardeau très différent.

L'instrument développé était le premier à évaluer le fardeau du traitement chez des patients souffrant de plusieurs maladies chroniques. Le développement et la validation de l'outil ont été réalisés en France. De part les relations entre fardeau et contexte de soin, il est probable que l'instrument aurait été différent s'il avait été développé dans un autre contexte de soin. Par exemple, dans notre étude, nous n'avons pas pris en compte les contraintes financières de la prise en charge médicale ; en effet, le système de sécurité sociale offre une prise en charge gratuite pour les patients souffrant de pathologies chroniques en France.

Pour mesurer le fardeau du traitement dans des contextes de soin différents de celui dans lequel l'outil a été développé, il faut donc adapter notre outil pour déterminer les

différents aspects culturels, du système de soins et de la prise en charge médicale pouvant avoir un impact sur la qualité de vie des patients.

3.7.4. Article

Tran VT, Montori VM, Eton DT, Baruch D, Falissard B, Ravaud P: Development and description of measurement properties of an instrument to assess Treatment Burden among patients with multiple chronic conditions. *BMC Med* 2012, 10(1):68

RESEARCH ARTICLE

Open Access

Development and description of measurement properties of an instrument to assess treatment burden among patients with multiple chronic conditions

Viet-Thi Tran^{1,2}, Victor M Montori³, David T Eton³, Dan Baruch⁴, Bruno Falissard^{5,6} and Philippe Ravaud^{1,2,7*}

Abstract

Background: Patients experience an increasing treatment burden related to everything they do to take care of their health: visits to the doctor, medical tests, treatment management and lifestyle changes. This treatment burden could affect treatment adherence, quality of life and outcomes. We aimed to develop and validate an instrument for measuring treatment burden for patients with multiple chronic conditions.

Methods: Items were derived from a literature review and qualitative semistructured interviews with patients. The instrument was then validated in a sample of patients with chronic conditions recruited in hospitals and general practitioner clinics in France. Factor analysis was used to examine the questionnaire structure. Construct validity was studied by the relationships between the instrument's global score, the Treatment Satisfaction Questionnaire for Medication (TSQM) scores and the complexity of treatment as assessed by patients and physicians. Agreement between patients and physicians was appraised. Reliability was determined by a test-retest method.

Results: A sample of 502 patients completed the Treatment Burden Questionnaire (TBQ), which consisted of 7 items (2 of which had 4 subitems) defined after 22 interviews with patients. The questionnaire showed a unidimensional structure. The Cronbach's α was 0.89. The instrument's global score was negatively correlated with TSQM scores ($r_s = -0.41$ to -0.53) and positively correlated with the complexity of treatment ($r_s = 0.16$ to 0.40). Agreement between patients and physicians ($n = 396$) was weak (intraclass correlation coefficient 0.38 (95% confidence interval 0.29 to 0.47)). Reliability of the retest ($n = 211$ patients) was 0.76 (0.67 to 0.83).

Conclusions: This study provides the first valid and reliable instrument assessing the treatment burden for patients across any disease or treatment context. This instrument could help in the development of treatment strategies that are both efficient and acceptable for patients.

Keywords: chronic disease/therapy, patient participation, physician-patient relations, quality of life, questionnaires, workload

Background

Chronic diseases are the leading cause of mortality in the world, representing more than 36 million deaths in 2008 [1]. About 45% of the population and 88% of people older than 65 years have at least one chronic condition. The prevalence of chronic diseases continues to increase: in 2020, nearly 50% of the US population will have at least one

chronic condition [2]. Therefore, the challenge for physicians has switched from curing acute illnesses to managing multiple chronic conditions. However, illnesses are still the primary focus of medical care [3] and many clinical practice guidelines focus on single conditions. For example, a physician following extant guidelines could prescribe up to 12 medications for a patient with osteoporosis, osteoarthritis, type 2 diabetes mellitus, hypertension, and chronic obstructive pulmonary disease [4].

* Correspondence: philippe.ravaud@htd.aphp.fr

¹Université Paris Descartes, Faculté de Médecine, Paris, France

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



Being a patient implies more investment of time and effort than just taking medicines. It also involves drug management, self-monitoring, visits to the doctor, laboratory tests and changes of lifestyle. For example, patients with type 2 diabetes controlled by oral agents could spend 143 minutes daily in recommended self-care [5]. This workload can affect quality of life as severely as the illness itself, and patients rate this treatment burden equal to that of diabetic neuropathy or nephropathy [6].

Treatment burden can be defined as the impact of health care on patients' functioning and well-being, apart from specific treatment side effects [7,8]. It takes into account everything patients do to take care of their health: visits to the doctor, medical tests, treatment management, and lifestyle changes. Treatment burden is associated, independently of illnesses, with adherence to therapeutic care [9,10] and could affect hospitalization [11] and survival rates [12].

Minimally disruptive medicine seeks to tailor treatment to the contexts of patients by integrating the notion of treatment burden in their care [13]. Therefore, caregivers need tools to establish the weight of the treatment burden. Many instruments assess treatment burden for specific conditions [14-18], but none has been developed to assess this burden globally across multiple chronic diseases. Because the treatment burden grows from the combination of chronic diseases, only an instrument that assesses it globally could help clinicians and researchers develop effective therapeutic programs that minimize the treatment workload [13].

In the present work, we aimed to develop a measure of treatment burden for patients with at least one chronic condition. This measure should be of use in daily clinical practice and in clinical research.

Methods

We used a multistep method to develop a tool to measure the treatment burden of chronic diseases [19,20] following the quality criteria proposed in the literature [21].

Stage 1: elaboration of the questionnaire

The objective of the instrument was to capture the perception of treatment burden of patients as 'the work of being a patient' dealing with increasingly complex treatment regimens [13], that is, the impact of the workload of healthcare on a patient's well-being and functioning.

We searched MEDLINE via PubMed for literature on treatment burden and existing questionnaires assessing it in specific diseases. We found no instrument appraising the treatment burden globally. Treatment burden was often assessed only as a subscale of specific disease scales [14-17] and thus was considered only for the regimen

associated with a particular condition. Items often focused on drug intake, adherence to care and convenience of use.

Using this literature review, three members of the team who had experience in the care of patients with chronic diseases (V-TT, BF, PR) highlighted possible relevant topics to capture the aspects of the workload of healthcare that could affect a patient's life. These topics were the burden associated with taking medicines, self-surveillance, laboratory tests, doctor visits, need for organization, administrative tasks, following advice on diet and physical exercise and social impact of the treatment. According to the conceptual model of our instrument, we chose not to include other consequences of the treatment such as treatment side effects.

In addition, because our instrument was elaborated in France and administered to French patients, we did not take into account the financial burden of treatment, because our national public health insurance program guarantees healthcare free of charge for patients with chronic conditions.

We recruited a convenience sample of 22 patients with at least 1 chronic condition from the department of internal medicine of Hospital Pitié-Salpêtrière and a general practitioner clinic in Paris in April 2011 (Additional file 1, Appendix 1). These two settings involved patients with various chronic conditions, requiring primary, secondary and tertiary care. During semistructured interviews, we presented the concept of treatment burden to patients and asked them about their diseases, their treatment and the burden of treatment, with open-ended questions: 'Could you tell us about your health problems?' 'Could you tell us about what you have to do to take care of your health?' 'What aspects of your care have the most impact on your life?' Then, we asked them about the burden associated with the different topics highlighted earlier by asking them (1) to rate each of these items, (2) to explain why they would rate it like that and (3) if they found the item relevant in the assessment of treatment burden generally. Finally, we asked patients, if other aspects of the workload of healthcare bothered them. As a result of these interviews, examples were added to the items, and we added one item 'Frequent healthcare reminds me of my health problems' to the questionnaire.

The resulting questionnaire consisted of seven items (two of which had four subitems), formed by an introductory sentence with examples, followed by a rating scale ranging from 0 to 10 with numbers placed under boxes and labeled end anchors ('No burden' and 'Considerable burden') [22-24].

A group of ten physicians (two methodologists, three general practitioners, two internists, one cardiologist, one pneumologist, one diabetologist) with experience in the care of patients with chronic conditions, some of whom

had experience in questionnaire development, reviewed the clarity and wording of the items. All physicians agreed that, on the surface, items appeared to be measuring what they actually were and that the instrument achieved face validity.

Stage 2: measurement properties of the instrument

The measurement properties of the questionnaire were assessed by four steps: (1) reduction of the number of items, (2) assessment of factorial validity, (3) assessment of construct validity and (4) assessment of reliability.

We recruited consecutive patients from six teaching hospitals of the Assistance-Publique Hôpitaux de Paris and eight general practitioner clinics in Paris to validate the questionnaire. Patients were eligible if they were 18 years or older, were able to complete a consent form and had at least one condition requiring medical follow-up for at least 6 months. Patients with cognitive impairment that could interfere with understanding the questionnaire were excluded. All patients provided written informed consent to be in the study.

Reducing the number of items was based on (1) a floor effect, considered present if more than 15% of respondents had the lowest score [21]; (2) the relevance of the items, assessed by the number of answers for which patients checked 'Does not apply'; and (3) item redundancy, suspected when interitem correlations by Spearman's correlation coefficient were > 0.80 [19]. Items were eliminated after discussion among three investigators (V-TT, BF, PR).

Answers to the questionnaire were aggregated in a global score by summing the item responses. 'Does not apply' or missing answers were considered the lowest possible score (0) because we considered that a patient not concerned by a domain of the treatment burden had no burden for that domain.

Factorial validity was assessed by determining the dimensional structure of the questionnaire by use of factor analysis. Scree plots were used to visualize a break between factors with large and small Eigenvalues. Factors that appeared before the horizontal break were assumed to be meaningful. Internal consistency was assessed by Cronbach's α [25] and was considered acceptable between 0.70 and 0.95 [26].

Construct validity was obtained by confirming two constructs theorized on the treatment burden [27]. First, we hypothesized a negative correlation between treatment burden, defined as the work of dealing with complex treatment regimens, and treatment satisfaction, defined as the balance between expectations about the treatment, side effects, convenience of use, and perceived efficacy. Treatment satisfaction was assessed by the Treatment Satisfaction Questionnaire for Medication (TSQM), an 11-item questionnaire validated in a population with diverse chronic conditions, measuring patient satisfaction with

various medications designed to treat, control or prevent a wide variety of medical conditions [28,29]. TSQM scores range from 0 to 100 and measure patient satisfaction with the treatment's effectiveness, side effects, convenience and globally. Correlations were expected to be higher between our instrument and the TSQM convenience score because some items overlapped. Second, we assumed a positive correlation between the patient evaluation of the treatment burden and treatment workload evaluated by items on (1) drug intake (number of tablets, injections and intakes per day); (2) medical follow-up (number of different physicians, medical appointments per month and hospitalizations per year); and (3) daily time spent on self-care. The correlations between the global questionnaire score, the TSQM scores and treatment workload variables were assessed by Spearman correlation coefficient (r_s) and considered high with $r_s > 0.50$ and moderate with r_s 0.35 to 0.50 [30]. Wilcoxon and Kruskal-Wallis tests were used to compare measurements for qualitative variables across groups. A P value < 0.05 was considered statistically significant. We used linear regression analyses to examine variables that predicted the global questionnaire score. Relationships were characterized with beta coefficients, standard errors, and percent variance explained (adjusted R^2) within these models. Heteroskedasticity was corrected by the method described by Greene *et al.* [31].

Description of our sample was completed by clustering homogenous groups of patients depending on the similarity of their response patterns to the Treatment burden questionnaire and analysis of treatment workload variables in each cluster of patients. Clustering involved a hierarchical ascendant classification with a Ward's distance method [32]. The number of clusters was determined so as to have a minimal sample of 100 patients. Stability of clustering was assessed by a twofold crossvalidation method.

We compared the patient's self-evaluation of treatment burden with an evaluation by their physician and by an informal caregiver using the same questionnaire adapted for heteroevaluation. Physicians and informal caregivers were asked to make the best estimate of the patient's treatment burden from their perspective.

Reliability of the instrument was determined by a test-retest method. Patients completed the new instrument twice: at baseline and after 2 weeks or 1 month. Reliability was assessed by the intraclass correlation coefficient (ICC) for agreement [33]. The 95% confidence intervals (95% CIs) were determined by a bootstrap method. Agreement was considered acceptable with $ICC > 0.60$ [27,34]. Agreement was represented by Bland and Altman plots, which represent the differences between two measurements against the means of the two measurements [35].

Statistical analyses involved use of SAS v. 9.2 (SAS Institute, Cary, NC, USA) and R v. 2.13.1 <http://www.r-project>.

org/. This study was approved by the Institutional Review Board of Hospital Bichat (IRB: 00006477).

Results

In total, 502 patients (mean age 59.3 (\pm 17) years; 266 women (53.1%)) were included to validate the questionnaire from April 2011 to September 2011 in France; 257

were inpatients (51.2%) and 300 reported a symptomatic disease (62.6%) (Table 1). Self-reported main chronic conditions ranged from diabetes (16.5%) to cancers (6.9%) and included well controlled psychiatric illnesses (1.6%).

During item reduction, we eliminated the subitem 'The conditions to store your medications (in your refrigerator etc.)' because a large number of patients responded 'Does

Table 1 Demographic and clinical characteristics of patients (n = 502)

Patient characteristics	Value	Missing data
Age, years (range)	60 (19 to 94)	2
Female sex, no. (%)	266 (53.1%)	1
Marital status, no. (%):		13
Married	216 (44.2%)	
Live-in partner	38 (7.8%)	
Single/separated	171 (35.0%)	
Widowed	64 (13.1%)	
Highest education level, no. (%):		44
No diploma/primary school	85 (18.6%)	
Secondary/high school	195 (42.6%)	
College	178 (38.9%)	
Inpatient, no. (%)	257 (51.2%)	
Duration of disease, years (range)	10 (0 to 91)	33
Presence of daily symptoms, no. (%)	300 (62.6%)	23
Need for assistance, no. (%)	132 (26.4%)	2
Hospitalizations during the last 12 months, no. (%)	0 (0 to 15)	40
Medical appointments/month, no. (%)	1 (0 to 30)	16
Different physicians, no. (%)	2 (0 to 10)	18
Tablets/day, no. (%)	4 (0 to 30)	14
Drug intakes/day, no. (%)	2 (0 to 6)	26
Injections/day, no. (%)	0 (0 to 8)	78
Diet, no. (%)	198 (40.3%)	11
Physical therapy, no. (%)	113 (22.9%)	9
Oxygen therapy, no. (%)	22 (4.4%)	4
Need for a specific organization for daily care, no. (%)	338 (67.3%)	
Time needed to organize drugs/week ^a	60 min (0 to 21 h)	
Need for self-monitoring, no. (%)	168 (33.47%)	
Time needed for self-monitoring/week ^a	60 min (0 to 12 h)	
Presence of side effects, no. (%)	168 (36.3%)	39
Main chronic condition, no. (%):		11
Diabetes	81 (16.5%)	
Rheumatologic diseases	59 (12.0%)	
High blood pressure and dyslipidemia	44 (9.0%)	
Systemic diseases	43 (8.8%)	
Pulmonary diseases (other than asthma)	40 (8.1%)	
Heart diseases	37 (7.5%)	
Asthma	37 (7.5%)	
Cancers and hematological malignancy	34 (6.9%)	
HIV infection	19 (3.9%)	
Arterial or venous thrombosis	17 (3.5%)	
Other diseases ^b	80 (16.3%)	

^aMedian time needed for concerned patients (self-reported)

^bOther diseases include diseases of the digestive system, psychiatric diseases, allergies, non-malignant hemopathy, neurological diseases, sequelae of injury, and endocrine diseases (other than diabetes).

not apply' (51.6%) and it had a large floor effect (64.0%) (Additional file 2, Appendix 2). Therefore, the final version of the questionnaire, the Treatment Burden Questionnaire (TBQ), consisted of seven items (two of which had four subitems) (Table 2).

Factorial validity, assessed by scree plots, favored a uni-dimensional instrument because 91% of the variance was explained by the first principal factor (Figure 1 and Additional file 3, Appendix 3). Cronbach's α was 0.89. The global score of the Treatment Burden Questionnaire was the sum of the answers to each item and ranged from 0 to 130. It was highly correlated with every item of the questionnaire ($r_s = 0.47$ to 0.68) (Additional file 4, Appendix 4).

Construct validity showed (1) a moderate negative correlation of the Treatment Burden Questionnaire score with the TSQM global and convenience scores ($r_s = -0.41$ and $r_s = -0.53$) and a weak negative correlation with the TSQM efficacy score ($r_s = -0.26$) (Table 3) and (2) a significant association of scores for variables used to describe treatment workload and the Treatment Burden Questionnaire global score (Table 4).

Using hierarchical ascendant classification, we clustered our sample in three homogenous groups of patients by the answers to the Treatment Burden Questionnaire (Additional file 5, Appendix 5). Twofold cross validation showed stable clustering results. The global score was $11.3 (\pm 9.2)$ in the first cluster, $34.6 (\pm 11.1)$ in the second cluster and $65.8 (\pm 18.1)$ in the third cluster. Therefore, we defined the clusters as patients with low, moderate and high burden of treatment. Descriptive analysis of the treatment workload items within the three clusters showed that

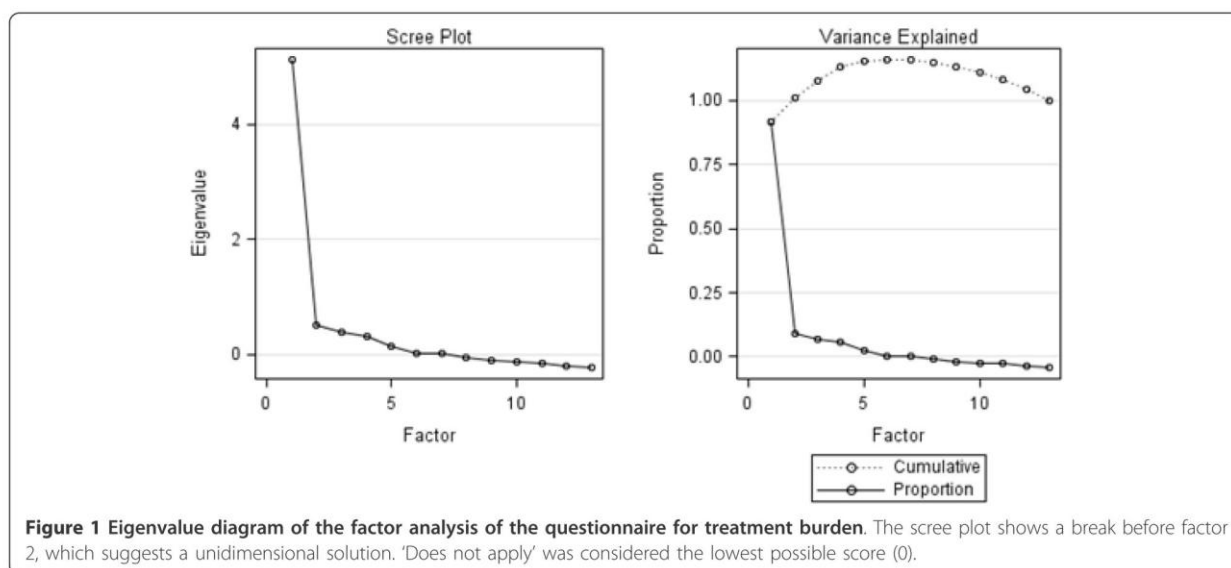
scores for these variables were significantly higher for patients with high treatment burden (Table 5). Treatment workload variables could explain up to 69% of the variability in the patient's score. Prediction of global score with these variables was more accurate with high than low treatment burden ($R^2 = 0.86$ vs $R^2 = 0.62$) (Additional file 6, Appendix 6). Treatment burden score was significantly higher when patients experienced medication side effects ($P < 0.0001$) and for patients whose treatment did not relieve their symptoms ($P < 0.0001$).

We found a moderate agreement (ICC 0.60 (0.28 to 0.79)) between patient and informal caregiver global scores (39 informal caregivers (7.8%) completed the questionnaire) (Additional file 7, Appendix 7a). Bland and Altman plots showed a mean difference of -8.7; 95% limits of agreement were -58.0 and 40.7 (Additional file 7, Appendix 7b). Agreement between patient and physician global scores was weak (ICC 0.38 (0.29 to 0.47)) (396 physicians (78.9%) completed the questionnaire) (Additional file 8, Appendix 8). Bland and Altman plot showed a mean difference of -7.6; 95% limits of agreement were -60.7 and 45.4 (Figure 2). Agreement between patient and general practitioner (n = 209) evaluations was ICC = 0.42 (0.27 to 0.54). Agreement between patient and hospital specialists (n = 187) evaluations was ICC = 0.29 (0.14 to 0.42) (Additional file 9, Appendix 9). Treatment workload variables could explain up to 76% of the variation in physician evaluations and was more accurate for patients with high than low treatment burden ($R^2 = 0.82$ vs $R^2 = 0.72$) (Additional file 6, Appendix 6).

Retests were obtained for 211 patients (42.0%). For the global score, the ICC for all retests was 0.76 (0.67 to 0.83)

Table 2 Items of the final Treatment Burden Questionnaire

Item no.	Item
1A	The taste, shape or size of your tablets and/or the inconvenience caused by your injections (for example, pain, bleeding, scars)
1B	The number of times you have to take your medication every day
1C	The things you do to remind yourself to take your daily medication and/or to manage your treatment when you are not at home
1D	The specific conditions when taking your medication (for example, taking it at a specific time of the day or meal, not being able to do certain things after taking them like driving or lying down)
2A	Lab tests and other exams (frequency, time spent and inconvenience of these exams)
2B	Self-monitoring (for example, taking your blood pressure or measuring your blood sugar yourself: frequency, time spent and inconvenience of this surveillance)
2C	Doctors visits (frequency and time spent for the visits)
2D	Arrange appointments and schedule doctors visits and lab tests
3	How would you rate the burden associated with taking care of paperwork from health insurance agencies, welfare organizations, hospitals and/or social care?
4	How would you rate the constraints associated with your diet (for example, not being allowed to eat certain foods)?
5	How would you rate the burden associated with the recommendations from your doctors to practice regular physical exercises?
6	What is the impact of your healthcare on your social relationships (for example, need for assistance, being ashamed to take your medication in front of people)?
7	'Frequent healthcare reminds me of my health problems'



(Additional file 10, Appendix 10a). Bland and Altman plots showed a mean difference of -5.9; 95% limits of agreement were -42.4 and 30.5 (Additional file 10, Appendix 10b). Reliability for the 2-week retest group (n = 182) was consistent with the 1-month retest group (n = 29) (ICC = 0.75 (0.65 to 0.83) vs ICC = 0.78 (0.46 to 0.91)).

Discussion

In this study, we presented a unidimensional valid and reliable instrument assessing the treatment burden of chronic diseases for patients with multiple chronic conditions. This patient-reported measure took into account the burden associated with drug intake, surveillance, lifestyle changes and the impact of healthcare on social relationships.

The instrument could help in clinical research for developing clinical practice guidelines adapted to the realities of patient lives. In addition, it could be used in clinical practice as a validated global score that is easy to calculate to identify patients overwhelmed by their treatment to help begin conversations about treatment burden with these patients.

We highlighted a negative correlation between treatment burden and treatment satisfaction: the more satisfied patients were with their treatment, the less the treatment burden. We expected that our scale score would correlate highly with the TSQM convenience score because some items overlapped. However, patients with side effects and who found the treatment inefficient would feel less agreeable to integrate the treatment in their lives.

Treatment burden did not concern only patients taking a lot of medications: 25% of patients in our sample took < 3 medications a day and still had a median treatment burden score of 17 (Q1 to Q3: 6 to 36). Therefore, treatment burden should be taken into account for every patient, because it could be associated with adherence to care [9] and thus could contribute to hospitalizations and survival rates [11]. However, physicians were often not fully aware of their patients' investment of time and efforts to comply with every prescription: we found only weak agreement between evaluation of treatment burden between patients and physicians. Even for specific domains such as self-monitoring or the prescription of a diet, physicians could not predict their patient's

Table 3 Relationship between the Treatment Satisfaction Questionnaire for Medication (TSQM) scores and Treatment Burden Questionnaire global score (n = 502 patients)

	Correlation with the Treatment Burden Questionnaire global score ^a	P value
TSQM global score	-0.41	< 0.0001
TSQM efficacy score	-0.26	< 0.0001
TSQM convenience score	-0.53	< 0.0001
TSQM side effects score ^a	-0.52	< 0.0001

The TSQM assesses satisfaction with medication. Scores range from 0 to 100. A high score indicates high satisfaction with the medication. Negative coefficients indicate a decrease in the TSQM score associated with an increase in treatment burden.

^aTSQM side effects score was calculated only for patients who declared experiencing side effects.

Table 4 Relationship between treatment workload variables and the Treatment Burden Questionnaire global score (n = 502 patients).

	Correlation with the Treatment Burden Questionnaire global score ^a	No. (%)	Mean score	P value
Correlation with treatment workload continuous variables				
Number of hospitalizations during the last year	0.24	-	-	< 0.0001
Number of medical appointments/month	0.28	-	-	< 0.0001
Number of different physicians	0.29	-	-	< 0.0001
Total number of tablets/day	0.25	-	-	< 0.0001
Total number of injections/day	0.31	-	-	< 0.0001
Number of drug intakes/day	0.16	-	-	0.0004
Time needed for healthcare per week (sum of the time needed for surveillance and the time needed to organize the treatment)	0.4	-	-	< 0.0001
Comparison between groups defined by treatment workload qualitative variables				
Need for a specific organization for daily care:				
Yes	-	338 (67.3%)	34.41	
No		164 (32.7%)	23.85	< 0.0001
Need for self-monitoring:				
Yes	-	168 (33.5%)	41.24	< 0.0001
No		334 (66.5%)	25.79	
Diet:				
Yes	-	198 (40.3%)	38.34	< 0.0001
No		293 (59.7%)	26.01	
Physical therapy:				
Yes	-	113 (22.9%)	35.77	
No		380 (77.1%)	29.73	0.02
Oxygen therapy:				
Yes	-	22 (4.4%)	44.68	
No		476 (94.8%)	30.47	0.005
Presence of side effects:				
Yes	-	168 (36.3%)	42.07	< 0.0001
No		295 (63.7%)	25.08	
The patient considers treatment as efficient:				
Yes	-	317 (63.1%)	27.35	< 0.0001
No		185 (36.8%)	37.13	

Spearman correlation coefficient for continuous variables and two-sided Wilcoxon two-sample test for qualitative variables

^aGlobal score is the sum of all items scores of the questionnaire with 'Does not apply' and missing answers considered as the lowest possible score (0).

evaluation. General practitioners, who are coordinators of care in France, have better knowledge than hospital specialists of how patients cope with everything they do to take care of their health (ICC = 0.42 for general practitioners and 0.20 for hospital specialists) but still fail to assess patients' treatment burden accurately. This finding is not unexpected, because treatment burden is a relatively new concept to physicians [13] and expresses a patient experience that is not shared in depth during consultations [36].

In existing questionnaires, treatment burden was often considered only as a subscale for larger disease-specific scales [16,17] and focused on a single treatment regimen. Given the increasing number of patients with multiple chronic diseases and complex treatment regimens, measuring global treatment burden seems increasingly

important. As Gallacher *et al.* have shown for chronic heart failure, treatment burden relates to how patients cope with their treatment [37]: (1) learning about treatment and their consequences, (2) monitoring the treatment, (3) adhering to treatment and lifestyle changes and (4) engaging with others. During our study, we asked patients about aspects of their healthcare that were not mentioned in our questionnaire but had an impact on their lives. We found the same domains of treatment burden as Gallacher *et al.*, with the exception of gaining an understanding about illness and treatments. Nevertheless, acquiring this knowledge is an important burden in the management of chronic conditions, especially when patients have to make sense of the disparate and conflicting information they gather from different sources. However, because we recruited patients with illnesses for at

Table 5 Characteristics of groups clustered by the hierarchical ascendant classification (n = 502 patients)

Characteristic	Whole sample (n = 502)	'Low burden' (n = 240)	'Moderate burden' (n = 140)	'High burden' (n = 122)	P value
Global score	30.1 ± 25.3	11.3 ± 9.2	34.6 ± 11.1	65.8 ± 18.1	< 0.0001
Age, years	59.3 ± 17.0	62.9 ± 16.1	59.2 ± 17.9	52.2 ± 15.6	< 0.0001
Female sex, no. (%)	266	118 (49.2%)	78 (55.7%)	70 (57.8%)	0.23
Marital status					0.19
Married	216	105 (45.3%)	60 (43.5%)	51 (42.9%)	
Live in partner	38	13 (5.6%)	12 (8.7%)	13 (10.9%)	
Single/separated	171	76 (32.8%)	49 (35.5%)	46 (38.7%)	
Widowed	64	38 (16.4%)	17 (12.3%)	9 (7.6%)	
Highest education level, no. (%)					0.76
No diploma/primary school	85	43 (19.9%)	23 (17.6%)	19 (17.1%)	
Secondary/high school	195	96 (44.4%)	54 (41.2%)	45 (40.5%)	
College	178	77 (35.6%)	54 (41.2%)	47 (42.3%)	
Inpatient, no. (percentage of the whole sample)	257	105 (43.7%)	84 (60.0%)	68 (55.7%)	0.004
Duration of disease, years	15.0 ± 15.4	16.3 ± 16.5	14.5 ± 16.7	13.3 ± 10.7	0.34
Presence of daily symptoms	300	110 (48.0%)	93 (70.4%)	97 (82.2%)	< 0.0001
Need for assistance, no. (%)	132	45 (18.8%)	41 (29.3%)	46 (38.0%)	0.0003
Number of hospitalizations during the last 12 months	1.1 ± 1.9	0.9 ± 1.8	1.2 ± 1.9	1.5 ± 2.1	0.0004
Number of medical appointments/month	2.3 ± 3.4	1.5 ± 1.6	2.6 ± 3.3	3.6 ± 5.2	< 0.0001
Number of different physicians	2.4 ± 1.4	2.1 ± 1.3	2.4 ± 1.3	2.9 ± 1.4	< 0.0001
Number of tablets/day	5.4 ± 4.5	4.5 ± 4.1	5.7 ± 3.9	7.0 ± 5.5	< 0.0001
Number of drug intakes/day	1.9 ± 1.0	1.8 ± 1.0	2.0 ± 1.0	2.1 ± 1.1	0.07
Number of injections/day	0.5 ± 1.3	0.2 ± 0.9	0.3 ± 0.9	1.3 ± 1.9	< 0.0001
Diet, no. (%)	198	75 (31.9%)	57 (42.2%)	66 (54.5%)	0.0002
Physical therapy, no. (%)	113	45 (19.2%)	33 (23.9%)	35 (28.9%)	0.11
Oxygen therapy, no. (%)	22	5 (2.1%)	8 (5.7%)	9 (7.4%)	0.05
Time needed to organize drugs/week*	22 ± 92 min	17 ± 100 min	13 ± 51 min	43 ± 108 min	< 0.0001
Need for self-monitoring, no. (%)	168	61 (25.4%)	36 (25.7%)	71 (58.2%)	< 0.0001
Time needed for self-monitoring/week*	14 ± 66 min	5 ± 31 min	10 ± 67 min	37 ± 101 min	< 0.0001
Presence of side effects, no. (%)	168	46 (20.9%)	56 (43.1%)	66 (58.4%)	< 0.0001
Patient considers his treatment efficient, no. (%)	317	176 (73.3%)	74 (52.9%)	67 (54.9%)	< 0.0001
Main chronic condition, no. (%)					< 0.0001
Diabetes	81	25 (10.7%)	14 (10.1%)	42 (35.0%)	
Rheumatologic diseases	59	25 (10.7%)	21 (15.2%)	13 (10.8%)	
Pulmonary diseases (other than asthma)	40	23 (9.9%)	12 (8.7%)	5 (4.2%)	
High blood pressure and dyslipidemia	44	29 (12.4%)	9 (6.5%)	6 (5.0%)	
Asthma	37	23 (9.9%)	8 (5.8%)	6 (5.0%)	
Systemic diseases	43	16 (6.9%)	13 (9.4%)	14 (11.7%)	

Patients were clustered in three groups depending on the similarity of their responses to the instrument. Global score was 11.3 (± 9.2) in the first cluster, 34.6 (± 11.1) in the second cluster and 65.8 (± 18.1) in the third cluster. Therefore, we defined the clusters as patients with low, moderate and high burden of treatment. Continuous variables are presented as mean ± SE. Categorical variables are presented as proportion of the corresponding subgroup. Associations between continuous variables among different classes were determined by Wilcoxon test. Qualitative variables are presented by their frequency in the whole sample. Associations between qualitative variables among different classes were determined by the χ^2 test. Global score is the sum of all items scores of the questionnaire with 'Does not apply' and missing answers considered the lowest possible score (0).

*Time needed for patients who did not require specific organization for daily care or who had no self-monitoring was considered 0.

least 6 months, they might have already coped with this particular burden, adapted to it, and therefore did not mention it.

The strengths of this study included field testing the instrument in a large sample of both inpatients and outpatients with different conditions and treatment regimens, which ensured that our instrument was flexible

enough for assessing the treatment burden across any disease or context. However, we found a significant floor effect and a large proportion of 'Does not apply' responses for all of our scales. This result was expected because treatment burden depends on how patients cope with their treatment regimens. Therefore, patients could have no burden in aspects of their care they have

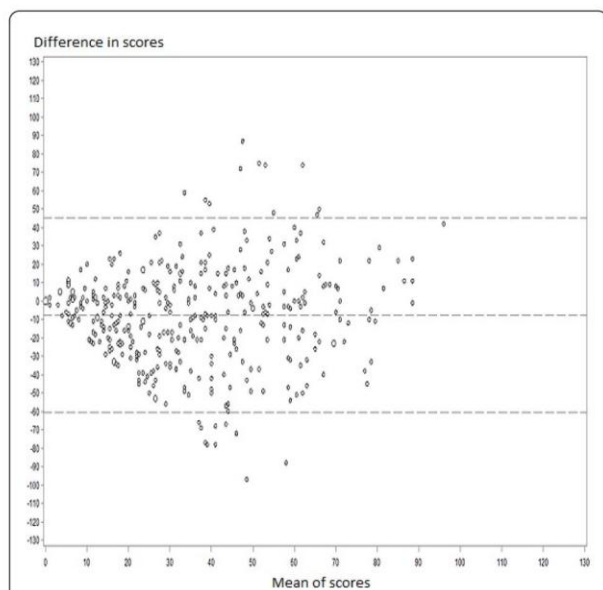


Figure 2 Bland and Altman plot representing agreement between Treatment Burden Questionnaire global scores for patients and physicians (n = 396 patients). The difference between global score for patients and physicians is plotted against the mean score. Negative differences mean that physicians overestimated the burden and positive differences that they underestimated it. Horizontal lines are drawn at the mean difference between the two measurements and the upper and lower limits of agreement. The size of markers reflects the number of individual observations.

integrated in their lives. As well, patients with similar treatment regimens could have very different treatment burdens. Still, domains not included in this instrument may be critical to some of these patients. During the validation study, we systematically searched for other aspects of treatment burden that could have an impact on patients' quality of life but found no preeminent domain.

More work in measuring treatment burden is needed. Because treatment burden depends on the context of patients (social or family structure, care delivery system) [13] and because our instrument was developed in France, we could not exclude that different domains could arise in other settings. As an example, the financial burden of the treatment did not arise from our qualitative interviews because the public health insurance program in France guarantees healthcare free of charge for patients with chronic conditions. In addition, depending on the social or family structure, the treatment burden may be shared by the patient with one or more informal caregivers, thus affecting the validity of the measure when only reported by the patient.

Conclusions

Our instrument on treatment burden for patients exhibiting multiple chronic conditions provides the first

valid and reliable solution to assess the burden of treatment across any disease or treatment context. It may help in the development of treatment strategies that are both efficient and acceptable for patients.

Additional material

Additional file 1: Appendix 1. Demographic and clinical characteristics of patients included in the semistructured interview pretest (n = 22).

Additional file 2: Appendix 2. Characteristics of the items presented to patients (n = 502 patients).

Additional file 3: Appendix 3. Eigenvalues for the correlation matrix.

Additional file 4: Appendix 4. Association of items of the Treatment Burden Questionnaire and global score (n = 502 patients).

Additional file 5: Appendix 5. Dendrogram of the hierarchical ascendant classification of patients by their answers to the Treatment Burden Questionnaire (n = 502 patients).

Additional file 6: Appendix 6. (a) Linear regression analysis of relation between global score for Treatment Burden Questionnaire as assessed by patients with variables associated with treatment workload. (b) Linear regression analysis of relation between global score for Treatment Burden Questionnaire as assessed by physicians with variables associated with treatment workload.

Additional file 7: Appendix 7. (a) Agreement between patients and informal caregiver evaluations of the treatment burden (n = 39). (b) Bland and Altman plot representing agreement between Treatment Burden Questionnaire global scores for patients and informal caregivers (n = 39).

Additional file 8: Appendix 8. Agreement between patients and physician evaluations of the treatment burden (n = 396).

Additional file 9: Appendix 9. Validation of the instrument in different subgroups.

Additional file 10: Appendix 10. (a) Reliability using test-retest (n = 211). (b) Bland and Altman plot representing the test-retest reliability of the Treatment Burden Questionnaire global score (n = 211).

Acknowledgements

The authors thank Michel Nougairède, Anna Christidis, Gaële Duriez-Mize, Marie-Anne Puel, Philippe Fabri, Ismael Nureni-Banafunzi, Elena Kisseleva, Gilbert Dhumerelle, Anne Simon, Serge Herson, Jacques Blacher, Christian Boitard, Gérard Huchon, Jean-François Bergmann, Alain Cohen-Solal, Olivier Meyer, Serge Poiraudou and Loïc Guillevin for permission to collect data; Clémence Palazzo for assistance in data acquisition; Gabriel Baron and Viet-Chi Tran for help in the analysis of the data; and Laura Smales for reviewing the manuscript before submission.

The TBQ[®] is protected by international copyright, with all rights reserved to PR and V-TT. For information on, or permission to use, the TBQ, please contact PROinformation@mapi-trust.org <http://www.mapi-trust.org>. This study was partly funded by INSERM U738, Paris, France.

Author details

¹Université Paris Descartes, Faculté de Médecine, Paris, France. ²INSERM U738, Paris, France. ³Division of Health Care and Policy Research, Department of Health Sciences Research and Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, USA. ⁴Université Paris Denis-Diderot, Faculté de Médecine, Paris, France. ⁵INSERM U669, Paris, France. ⁶Université Paris Sud, Paris, France. ⁷Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA.

Authors' contributions

V-TT, BF and PR conceived and designed the study. V-TT and DB acquired the data. V-TT, BF, PR analyzed and interpreted the data. V-TT and PR drafted the manuscript. VM, BF, DE, DB and PR critically revised the manuscript for

important intellectual content. DB and PR provided administrative, technical, and material support. All authors saw and approved the final manuscript. PR is the guarantor, had full access to the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests

The authors declare that they have no competing interests.

Received: 13 January 2012 Accepted: 4 July 2012 Published: 4 July 2012

References

1. Alwan A: *Global Status Report on Noncommunicable Diseases 2010* Geneva, Switzerland: World Health Organization; 2011.
2. Wolff JL, Starfield B, Anderson G: **Prevalence, expenditures, and complications of multiple chronic conditions in the elderly.** *Arch Intern Med* 2002, **162**:2269-2276.
3. Tinetti ME, Fried T: **The end of the disease era.** *Am J Med* 2004, **116**:179-185.
4. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW: **Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance.** *JAMA* 2005, **294**:716-724.
5. Russell LB, Suh DC, Safford MA: **Time requirements for diabetes self-management: too much for many?** *J Fam Pract* 2005, **54**:52-56.
6. Huang ES, Brown SE, Ewigman BG, Foley EC, Meltzer DO: **Patient perceptions of quality of life with diabetes-related complications and treatments.** *Diabetes Care* 2007, **30**:2478-2483.
7. Eton DT, Mullan RJ, Yost KJ, Egginton J, Ridgeway J, Fields HE, Mustafa M, Schmit K, Erwin PJ, Murad H, Montori VM: **A systematic review of burden of treatment instruments in three chronic diseases.** *Int Soc Qual Life Res* 2011, **86**(Suppl), ab. 294.
8. Eton DT, Oliveira DR, Egginton J, Mair FS, May C, Montori VM: **Understanding the burden of treatment in patients with multiple chronic conditions: Evidence from exploratory interviews.** *Qual Life Res* 2010, **19**(Suppl 1), ab. 1673.
9. Vijan S, Hayward RA, Ronis DL, Hofer TP: **Brief report: the burden of diabetes therapy: implications for the design of effective patient-centered treatment regimens.** *J Gen Intern Med* 2005, **20**:479-482.
10. Vermeire E, Hearnshaw H, Van Royen P, Denekens J: **Patient adherence to treatment: three decades of research. A comprehensive review.** *J Clin Pharm Ther* 2001, **26**:331-342.
11. Ho PM, Rumsfeld JS, Masoudi FA, McClure DL, Plomondon ME, Steiner JF, Magid DJ: **Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus.** *Arch Intern Med* 2006, **166**:1836-1841.
12. Rasmussen JN, Chong A, Alter DA: **Relationship between adherence to evidence-based pharmacotherapy and long-term mortality after acute myocardial infarction.** *JAMA* 2007, **297**:177-186.
13. May C, Montori VM, Mair FS: **We need minimally disruptive medicine.** *BMJ* 2009, **339**:b2803.
14. Wilcox AR, Dragnev MC, Darcey CJ, Siegel CA: **A new tool to measure the burden of Crohn's disease and its treatment: do patient and physician perceptions match?** *Inflamm Bowel Dis* 2010, **16**:645-650.
15. Liu JY, Woloshin S, Laycock WS, Rothstein RI, Finlayson SR, Schwartz LM: **Symptoms and treatment burden of gastroesophageal reflux disease: validating the GERD assessment scales.** *Arch Intern Med* 2004, **164**:2058-2064.
16. Brod M, Hammer M, Kragh N, Lessard S, Bushnell DM: **Development and validation of the Treatment Related Impact Measure of Weight (TRIM-Weight).** *Health Qual Life Outcomes* 2010, **8**:19.
17. Brod M, Hammer M, Christensen T, Lessard S, Bushnell DM: **Understanding and assessing the impact of treatment in diabetes: the Treatment-Related Impact Measures for Diabetes and Devices (TRIM-Diabetes and TRIM-Diabetes Device).** *Health Qual Life Outcomes* 2009, **7**:83.
18. Prins MH, Guillemin I, Gilet H, Gabriel S, Essers B, Raskob G, Kahn SR: **Scoring and psychometric validation of the Perception of Anticoagulant Treatment Questionnaire (PACT-Q).** *Health Qual Life Outcomes* 2009, **7**:30.
19. Falissard B: *Mesurer la subjectivité en santé. Perspective méthodologique et statistique* Paris, France: Masson; 2001.
20. Guyatt GH, Bombardier C, Tugwell PX: **Measuring disease-specific quality of life in clinical trials.** *CMAJ* 1986, **134**:889-895.
21. Terwee CB, Bot SD, De Boer MR, Van der Windt DA, Knol DL, Dekker J, Bouter LM, De Vet HC: **Quality criteria were proposed for measurement properties of health status questionnaires.** *J Clin Epidemiol* 2007, **60**:34-42.
22. Frisbie D, Brandenburg D: **Equivalence of questionnaire items with varying response formats.** *J Educ Meas* 1979, **16**:43-48.
23. Wildt A, Mazis A: **Determinant of scale response: label versus position.** *J Marketing Res* 1978, **15**:261-267.
24. Schwarz N, Knauper B, Hippler H, Noelle-Neumann E, Clark L: **Rating scales: Numeric values may change the meaning of scale labels.** *Public Opin Quart* 1991, **55**:570-582.
25. Cronbach L: **Coefficient alpha and the internal structure of tests.** *Psychometrika* 1951, **16**:297-334.
26. Nunnally J, Bernstein I: *Psychometric Theory* New York, USA: McGraw-Hill; 1994.
27. Streiner D, Norman G: *Health Measurement Scales: A Practical Guide to Their Development and Use* Oxford, UK: Oxford University Press; 2008.
28. Atkinson MJ, Sinha A, Hass SL, Colman SS, Kumar RN, Brod M, Rowland CR: **Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease.** *Health Qual Life Outcomes* 2004, **2**:12.
29. Atkinson MJ, Kumar R, Cappelleri JC, Hass SL: **Hierarchical construct validity of the treatment satisfaction questionnaire for medication (TSQM version II) among outpatient pharmacy consumers.** *Value Health* 2005, **8**(Suppl 1):S9-S24.
30. Juniper E, Guyatt G, Jaeschke R: **How to develop and validate a new health-related quality of life instrument.** *Quality of Life and Pharmacoeconomics in Clinical Trials* New York, USA: Raven Press; 1996, 49-56.
31. Greene W: *Econometric Analysis* New York, USA: Macmillan; 2000.
32. Joe H, Ward J: **Hierarchical grouping to optimize an objective function.** *J Am Stat Assoc* 1963, **58**:236-244.
33. Shrout PE, Fleiss JL: **Intraclass correlations: uses in assessing rater reliability.** *Psychol Bull* 1979, **86**:420-428.
34. Lohr KN: **Assessing health status and quality-of-life instruments: attributes and review criteria.** *Qual Life Res* 2002, **11**:193-205.
35. Bland JM, Altman DG: **Measuring agreement in method comparison studies.** *Stat Methods Med Res* 1999, **8**:135-160.
36. Bohlen K, Scoville E, Shippee ND, May CR, Montori VM: **Overwhelmed patients: a videographic analysis of how patients with type 2 diabetes and clinicians articulate and address treatment burden during clinical encounters.** *Diabetes Care* 2012, **35**:47-49.
37. Gallacher K, May CR, Montori VM, Mair FS: **Understanding patients' experiences of treatment burden in chronic heart failure using normalization process theory.** *Ann Fam Med* 2011, **9**:235-243.

Pre-publication history

The pre-publication history for this paper can be accessed here:
<http://www.biomedcentral.com/1741-7015/10/68/prepub>

doi:10.1186/1741-7015-10-68

Cite this article as: Tran et al.: Development and description of measurement properties of an instrument to assess treatment burden among patients with multiple chronic conditions. *BMC Medicine* 2012 **10**:68.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Characteristic	Value
Age – yr	70 (53-76)
Female sex – no. (%)	11 (50.0%)
Marital status – no. (%)	
Married	11 (50.0 %)
Live in partner	2 (9.1 %)
Single/separated	3 (13.6 %)
Widowed	6 (27.3 %)
Highest education level –no. (%)	
No diploma /Primary school	7 (33.3 %)
Secondary / High school	4 (19.0 %)
College	10 (47.7 %)
Inpatient – no. (%)	6 (27.3 %)
Duration of disease – yr	11 (5-23)
Presence of daily symptoms – no. (%)	18 (85.7 %)
Need of assistance – no. (%)	7 (31.8 %)
Number of hospitalization during the last 12 months	0 (0-3)
Number of medical appointments/month	1 (0-3)
Number of different physicians	3 (2-4)
Number of tablets/day	5 (4-6)
Number of drug intakes/day	2 (2-2)
Number of injections/day	0 (0-0)
Diet – no. (%)	13 (59.1 %)
Physical therapy – no. (%)	6 (27.1 %)
Oxygen therapy – no. (%)	5 (22.0 %)

Appendix 1. Demographic and clinical Characteristics of patients included in the pretest (n=22).

Median (range) were reported for continuous variables and number (%) for categorical variables.

Item	Response rate No. (%)	Item “Does not apply” No. (%)	Floor effect No. (%)	“High burden” (score>7) No. (%)
1A. The taste, shape or size of your tablets and/or the inconvenience caused by your injections (e.g., pain, bleeding, scars)	496 (98.8 %)	91 (18.3 %)	180 (44.4 %)	44 (10.9 %)
1B. The number of times you have to take your medication every day	498 (99.2 %)	61 (12.2 %)	166 (38.0 %)	36 (8.2 %)
1C. The things you do to remind yourself to take your daily medication and/or to manage your treatment when you are not at home.	499 (99.4 %)	71 (14.2 %)	149 (34.8 %)	60 (14.0 %)
1D. The specific conditions when taking your medication (e.g., taking it at a specific time of the day or meal, not being able to do certain things after taking them like driving or lying down)	501 (99.8 %)	132 (26.3 %)	166 (45.0 %)	45 (12.2 %)
1E. The conditions for storing your medications (e.g., in your refrigerator)	500 (99.6 %)	258 (51.6 %)	155 (64.0 %)	29 (12.0 %)
2A. Lab tests and other exams (frequency, time spent and inconvenience of these exams)	500 (99.6 %)	31 (6.2 %)	104 (22.2 %)	85 (18.1 %)
2B. Self-monitoring (e.g., taking your blood pressure or measuring your blood sugar yourself: frequency, time spent and inconvenience of this surveillance)	499 (99.4 %)	243 (48.7 %)	91 (35.5 %)	43 (16.8 %)
2C. Doctors visits (frequency and time spent for the visits)	496 (98.8 %)	27 (5.4 %)	126 (26.9 %)	75 (16.0 %)
2D. Arrange appointments and schedule doctors visits and lab tests	499 (99.4 %)	35 (7.0 %)	129 (27.8 %)	87 (18.7 %)
3. How would you rate the burden associated with taking care of paperwork from health insurance agencies, welfare organizations, hospitals and/or social care?	499 (99.4 %)	56 (11.2 %)	146 (33.0 %)	84 (19.0 %)
4. How would you rate the constraints associated with your diet (e.g., not being allowed to eat certain food)?	500 (99.6 %)	183 (36.6 %)	87 (27.4 %)	73 (23.0 %)
5. How would you rate the burden associated with the recommendations from your doctors to practice regular physical exercises?	496 (98.8 %)	138 (27.8 %)	128 (35.7 %)	73 (20.4 %)
6. What is the impact of your healthcare on your social relationships (e.g., need for assistance, being ashamed to take your medication in front of people)?	498 (99.2 %)	84 (16.9 %)	218 (52.7 %)	67 (16.2 %)
7. "Frequent healthcare reminds me of my health problems"	495 (98.6 %)	-	187 (37.8 %)	128 (22.2 %)

Appendix 2: Characteristics of the items presented to patients (n=502 patients). Floor-effect frequencies and “high burden patients” frequencies relate to patients concerned by the item and thus not checking “Does not apply”.

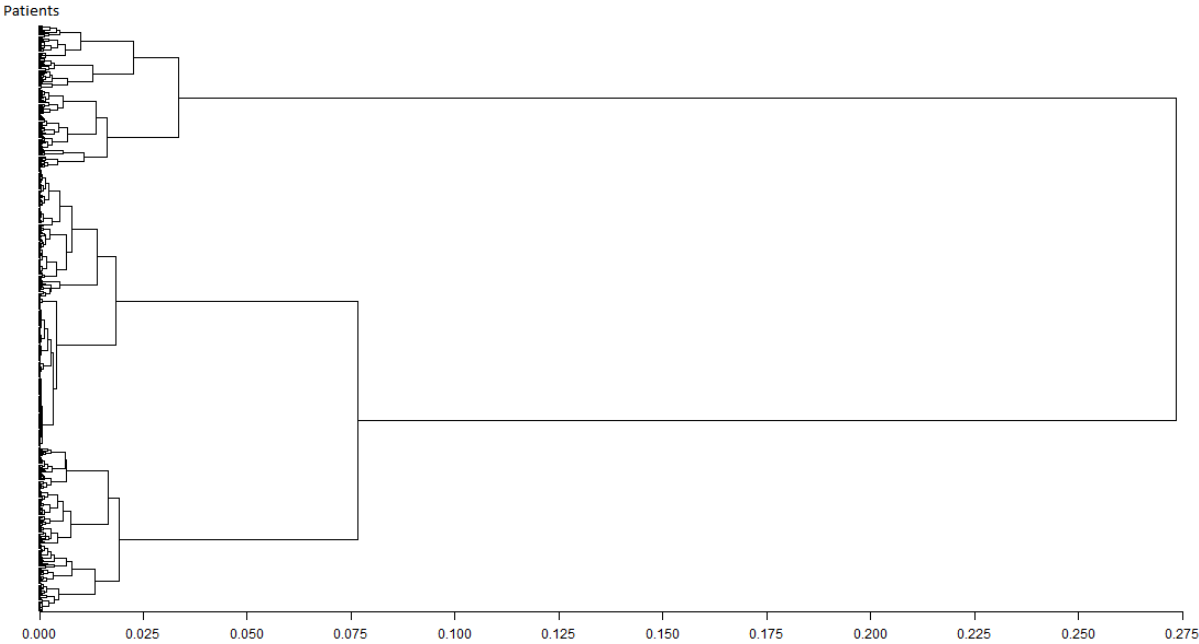
Appendix 3: Eigenvalues for the correlation matrix. “Does not apply” were considered as the lowest possible score (0).

	Eigenvalue	Proportion of variance	Cumulative variance
1	5.12	0.92	0.92
2	0.50	0.09	1.00

Appendix 4: Association of items of the Treatment Burden Questionnaire and global score (n=502 patients). Correlation of each individual item with the scale total, omitting that item from the total, by Spearman correlation coefficient.

Item	Spearman correlation coefficient
1A. The taste, shape or size of your tablets and/or the inconvenience caused by your injections (e.g., pain, bleeding, scars)	0.60 (p<0.0001)
1B. The number of times you have to take your medication every day	0.61 (p<0.0001)
1C. The things you do to remind yourself to take your daily medication and/or to manage your treatment when you are not at home.	0.54 (p<0.0001)
1D. The specific conditions when taking your medication (e.g., taking it at a specific time of the day or meal, not being able to do certain things after taking them like driving or lying down)	0.58 (p<0.0001)
2A. Lab tests and other exams (frequency, time spent and inconvenience of these exams)	0.68 (p<0.0001)
2B. Self-monitoring (e.g., taking your blood pressure or measuring your blood sugar yourself: frequency, time spent and inconvenience of this surveillance)	0.54 (p<0.0001)
2C. Doctors visits (frequency and time spent for the visits)	0.67 (p<0.0001)
2D. Arrange appointments and schedule doctors visits and lab tests	0.68 (p<0.0001)
3. How would you rate the burden associated with taking care of paperwork from health insurance agencies, welfare organizations, hospitals and/or social care?	0.55 (p<0.0001)
4. How would you rate the constraints associated with your diet (e.g., not being allowed to eat certain food)?	0.47 (p<0.0001)
5. How would you rate the burden associated with the recommendations from your doctors to practice regular physical exercises?	0.47 (p<0.0001)
6. What is the impact of your healthcare on your social relationships (e.g., need for assistance, being ashamed to take your medication in front of people)?	0.52 (p<0.0001)
7. "Frequent healthcare reminds me of my health problems"	0.57 (p<0.0001)

Appendix 5: Dendrogram of the hierarchical ascendant classification of patients by their answers to the Treatment Burden Questionnaire (n=502 patients). Hierarchical ascendant classification was performed using a Ward's distance method. Classification resulted in the formation of 3 clusters of patients which could be described as patients with low burden of treatment, medium burden of treatment and high burden of treatment.



Appendix 6(a): Linear regression analysis of relation between global score for Treatment Burden Questionnaire as assessed by patients with variables associated with treatment workload

Treatment workload variables significantly associated with global score in the whole sample (n=502 patients). $R^2=0.69$

	β	SE	p
Number of hospitalizations	1.69	0.68	0.01
Number of different physicians	4.19	0.68	<0.0001
Number of medical appointments/month	1.92	0.31	<0.0001
Number of tablets/day	1.28	0.28	<0.0001
Number of injections/day	5.30	1.21	<0.0001
Diet	6.05	2.23	0.007
Need of self monitoring	6.21	2.73	0.02

Treatment workload variables significantly associated with global score for patients with low burden of treatment as defined by the hierarchical ascendant classification (HAC) (n=240 patients). $R^2=0.62$

	β	SE	P
Number of different physicians	2.78	0.39	<0.0001
Number of medical appointments/month	1.90	0.40	<0.0001
Diet	4.97	1.34	0.0003

Treatment workload variables significantly associated with global score for patients with moderate burden of treatment as defined by the hierarchical ascendant classification (HAC) (n=140 patients). $R^2=0.77$

	β	SE	P
Number of hospitalizations	2.19	0.94	0.02
Number of different physicians	5.96	0.92	<0.0001
Number of tablets/day	1.72	0.39	<0.0001
Diet	9.48	3.23	0.004

Treatment workload variables significantly associated with the global score for patients with high burden of treatment as defined by the hierarchical ascendant classification (HAC) (n=122 patients). $R^2=0.86$

	β	SE	p
Number of different physicians	11.68	1.46	<0.0001
Number of tablets/day	2.32	0.55	<0.0001
Number of injections/day	4.36	1.89	0.02
Need of self monitoring	12.93	5.43	0.02

Appendix 6(b): Linear regression analysis of relation between global score for Treatment Burden Questionnaire as assessed by physicians with variables associated with treatment workload

Treatment workload variables significantly associated with the physician's global score in the whole sample (n=502). R²=0.76

	β	SE	p
Number of different physicians	5.57	0.84	<0.0001
Number of tablets/day	1.79	0.28	<0.0001
Number of injections/day	4.14	1.26	0.001
Diet	9.21	3.00	0.002
Need for self-monitoring	15.39	3.25	<0.0001

Treatment workload variables significantly associated with the physician's global score for patients with a low burden of treatment as defined by the hierarchical ascendant classification (HAC) (n=240). R²=0.72

	β	SE	p
Number of different physicians	5.53	1.16	<0.0001
Number of tablets/day	1.42	0.39	0.0005
Number of injections/day	4.81	2.41	0.05
Diet	9.43	4.43	0.03
Prescription of oxygen	24.37	10.42	0.02
Need for self-monitoring	13.40	4.71	0.005

Treatment workload variables significantly associated with the physician's global score for patients with a moderate burden of treatment as defined by the hierarchical ascendant classification (HAC) (n=140). R²=0.74

	β	SE	p
Number of different physicians	3.84	1.69	0.03
Number of tablets/day	2.72	0.61	<0.0001
Number of injections/day	8.15	3.21	0.01
Diet	14.77	5.58	0.01

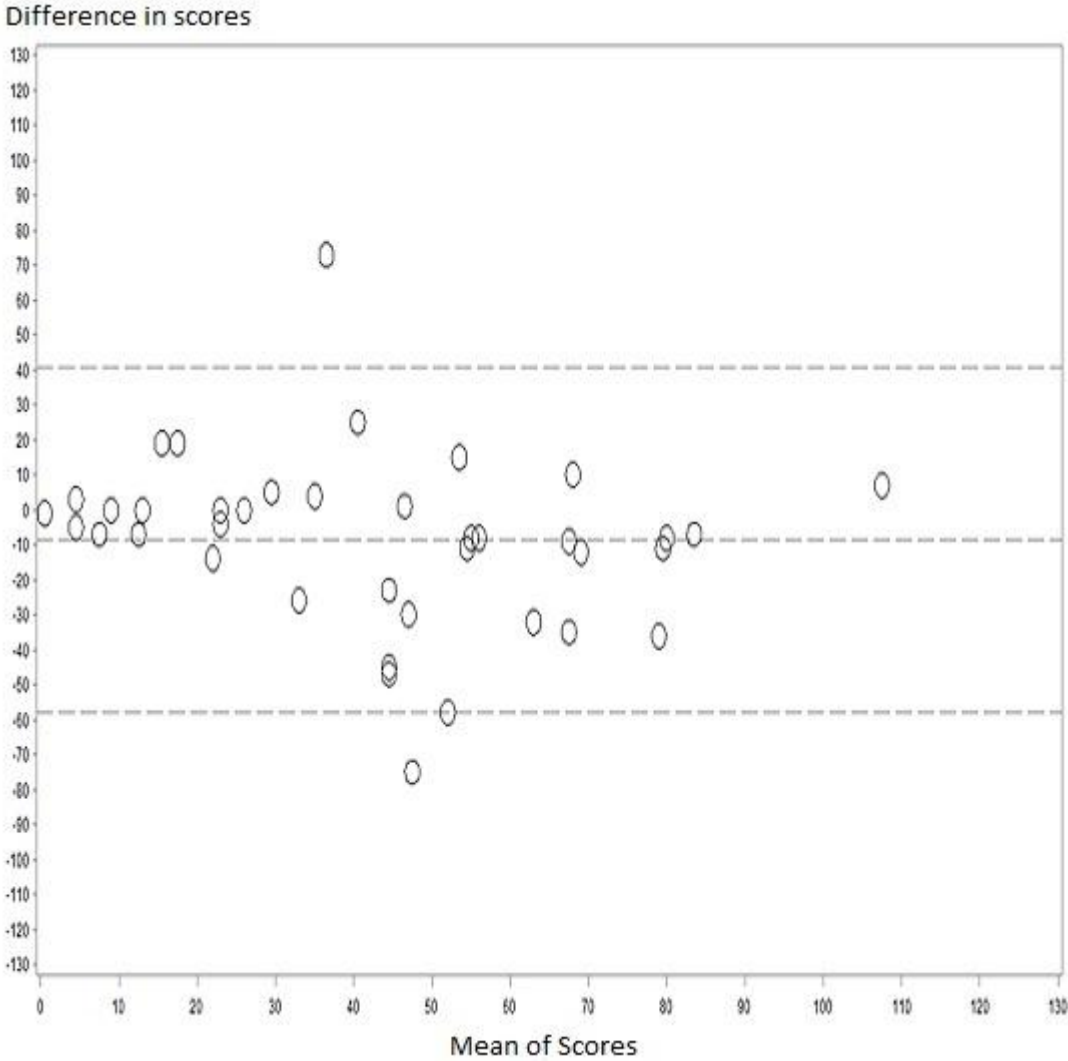
Treatment workload variables significantly associated with the physician's global score for patients with a high burden of treatment as defined by the hierarchical ascendant classification (HAC) (n=122). R²=0.82

	β	SE	p
Number of different physicians	8.67	1.51	<0.0001
Number of tablets/day	1.16	0.51	0.03
Need for self-monitoring	26.58	5.14	<0.0001

Item	ICC	95% CI
1A. The taste, shape or size of your tablets and/or the inconvenience caused by your injections (e.g., pain, bleeding, scars)	0.28	[0-0.60]
1B. The number of times you have to take your medication every day	0.58	[0.25-0.80]
1C. The things you do to remind yourself to take your daily medication and/or to manage your treatment when you are not at home.	0.59	[0.34-0.80]
1D. The specific conditions when taking your medication (e.g., taking it at a specific time of the day or meal, not being able to do certain things after taking them like driving or lying down)	0.50	[0.16-0.78]
2A. Lab tests and other exams (frequency, time spent and inconvenience of these exams)	0.41	[0-0.70]
2B. Self-monitoring (e.g., taking your blood pressure or measuring your blood sugar yourself: frequency, time spent and inconvenience of this surveillance)	0.46	[0.02-0.77]
2C. Doctors visits (frequency and time spent for the visits)	0.29	[0-0.60]
2D. Arrange appointments and schedule doctors visits and lab tests	0.44	[0.08-0.69]
3. How would you rate the burden associated with taking care of paperwork from health insurance agencies, welfare organizations, hospitals and/or social care?	0.47	[0.17-0.75]
4. How would you rate the constraints associated with your diet (e.g., not being allowed to eat certain food)?	0.53	[0.15-0.82]
5. How would you rate the burden associated with the recommendations from your doctors to practice regular physical exercises?	0.47	[0.09-0.76]
6. What is the impact of your healthcare on your social relationships (e.g., need for assistance, being ashamed to take your medication in front of people)?	0.41	[0.13-0.70]
7. "Frequent healthcare reminds me of my health problems"	0.48	[0.17-0.72]
Score global	0.58	[0.21-0.80]

Appendix 7(a): Agreement between patients and informal caregiver evaluations of the treatment burden (n=39) with ICC for agreement. 95% confidence intervals (95% CIs) were calculated by a bootstrap method. *Global score is the sum of all item scores with “Does not apply” and missing answers having the lowest possible score (0).

Appendix 7(b): Bland and Altman plot representing agreement between Treatment Burden Questionnaire global scores for patients and informal caregivers (n=39). The difference between patient and informal caregiver scores is plotted against the mean score. Negative differences indicate that informal caregivers overestimated the burden and positive differences that they underestimated it. Horizontal lines are drawn for the mean difference between the two measurements and the upper and lower limits of agreement.



Item	ICC	95% CI
1A. The taste, shape or size of your tablets and/or the inconvenience caused by your injections (e.g., pain, bleeding, scars)	0.23	[0.13-0.33]
1B. The number of times you have to take your medication every day	0.25	[0.15-0.35]
1C. The things you do to remind yourself to take your daily medication and/or to manage your treatment when you are not at home.	0.23	[0.14-0.33]
1D. The specific conditions when taking your medication (e.g., taking it at a specific time of the day or meal, not being able to do certain things after taking them like driving or to lying down)	0.22	[0.12-0.34]
2A. Lab tests and other exams (frequency, time spent and inconvenience of these exams)	0.30	[0.19-0.39]
2B. Self-monitoring (e.g., taking your blood pressure or measuring your blood sugar yourself: frequency, time spent and inconvenience of this surveillance)	0.41	[0.30-0.51]
2C. Doctors visits (frequency and time spent for the visits)	0.24	[0.15-0.34]
2D. Arrange appointments and schedule doctors visits and lab tests	0.22	[0.12-0.31]
3. How would you rate the burden associated with taking care of paperwork from health insurance agencies, welfare organizations, hospitals and/or social care?	0.28	[0.18-0.38]
4. How would you rate the constraints associated with your diet (e.g., not being allowed to eat certain food)?	0.38	[0.28-0.47]
5. How would you rate the burden associated with the recommendations from your doctors to practice regular physical exercises?	0.19	[0.09-0.29]
6. What is the impact of your healthcare on your social relationships (e.g., need for assistance, being ashamed to take your medication in front of people)?	0.20	[0.11-0.31]
7. "Frequent healthcare reminds me of my health problems"	0.21	[0.11-0.29]
Global score*	0.38	[0.29-0.47]

Appendix 8: Agreement between patients and physician evaluations of the treatment burden (n=396) with ICC for agreement. 95% CIs were calculated by a bootstrap method. *Global score is the sum of all item scores with “Does not apply” and missing answers having the lowest possible score (0).

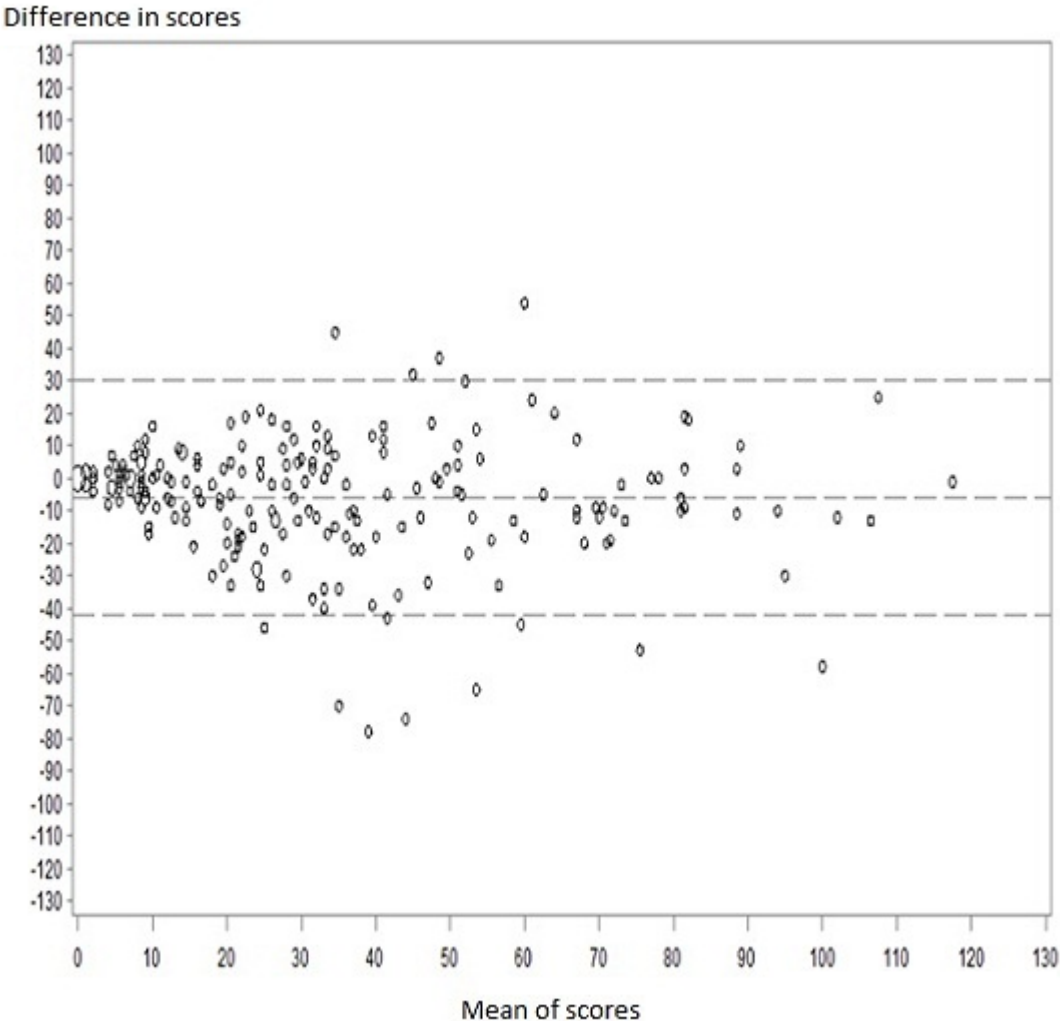
	Global score	p	Agreement between the physician and the patient	Correlation with TSQM scores using Spearman correlation coefficient	
				TSQM global score	TSQM convenience score
Age					
<60 (n= 243)	38.4 (± 26.7)	<0.0001	0.36 [0.25-0.49]	-0.39 (p<0.0001)	-0.53 (p<0.0001)
>60 (n= 259)	24.0 (± 21.9)		0.37 [0.23-0.47]	-0.36 (p<0.0001)	-0.47 (p<0.0001)
Presence of symptoms					
Symptomatic (n= 300)	37.6 (± 26.0)	<0.0001	0.41 [0.28-0.51]	-0.32 (p<0.0001)	-0.46 (p<0.0001)
Asymptomatic (n= 179)	20.6 (± 20.7)		0.27 [0.14-0.42]	-0.34 (p<0.0001)	-0.50 (p<0.0001)
Educational level					
No diploma/ Primary (n= 85)	29.0 (± 26.9)	0.04	0.54 [0.17-0.69]	-0.37 (p=0.0008)	-0.56 (p<0.0001)
Secondary/ high school (n= 195)	29.1 (± 24.5)		0.35 [0.21-0.48]	-0.44 (p<0.0001)	-0.50 (p<0.0001)
College (n= 178)	34.3 (± 25.3)		0.36 [0.18-0.51]	-0.41 (p<0.0001)	-0.57 (p<0.0001)
Inpatients (n=257)	34.7(± 27.7)	0.001	0.29 [0.14-0.42]	-0.36 (p<0.0001)	-0.50 (p<0.0001)
Outpatients (n= 245)	27.1(± 23.1)		0.42 [0.27-0.54]	-0.45 (<0.0001)	-0.56 (p<0.0001)
Different conditions reported as “main chronic condition”					
Diabetes (n=81)	46.4 (± 28.6)	<0.0001	0.23 [0-0.44]	-0.28 (p=0.01)	-0.55 (p<0.0001)
Rheumatologic diseases (n=59)	28.6 (± 26.3)		0.08 [0-0.42]	-0.41 (p=0.001)	-0.49 (p<0.0001)
High blood pressure and dyslipidemia (n=44)	18.5 (± 19.9)		0.38 [0.12-0.74]	-0.32 (p=0.04)	-0.38 (p=0.01)
Systemic diseases (n=43)	39.0 (± 26.3)		0.15 [0-0.46]	-0.34 (p=0.02)	-0.35 (p=0.02)
Pulmonary diseases (other than asthma) (n=40)	24.8 (± 17.5)		0.40 [0.10-0.64]	-0.55 (p=0.0003)	-0.56 (p=0.0003)
Heart diseases (n=37)	29.3 (± 23.7)		0.39 [0.07-0.67]	-0.17 (p=0.33)	-0.55 (p=0.0005)

Appendix 9: Validation of the instrument in different subgroups. Global score is presented as mean (± SD). Wilcoxon and Kruskal Wallis tests were used for comparing subgroups. Agreement between physician and patient measurements, with ICC for agreement and a bootstrap method for calculating 95% CIs, was consistent with results obtained for the whole sample except for rheumatologic diseases, where it was weaker. Spearman correlation coefficient was used to assess the correlations between the global score and the TSQM scores.

Item	All retests (n=211)	2-week interval retest (n=182)	1-month interval retest (n=29)
1A. The taste, shape or size of your tablets and/or the inconvenience caused by your injections (e.g., pain, bleeding, scars)	0.56 [0.43-0.69]	0.58 [0.44-0.71]	0.50 [0.09-0.82]
1B. The number of times you have to take your medication every day	0.53 [0.39-0.64]	0.57 [0.45-0.69]	0.35 [0-0.71]
1C. The things you do to remind yourself to take your daily medication and/or to manage your treatment when you are not at home.	0.57 [0.44-0.68]	0.56 [0.42-0.68]	0.61 [0.30-0.87]
1D. The specific conditions when taking your medication (e.g., taking it at a specific time of the day or meal, not being able to do certain things after taking them like driving or lying down)	0.49 [0.35-0.61]	0.48 [0.31-0.62]	0.52 [0.21-0.79]
2A. Lab tests and other exams (frequency, time spent and inconvenience of these exams)	0.62 [0.50-0.72]	0.61 [0.45-0.71]	0.71 [0.47-0.90]
2B. Self-monitoring (e.g., taking your blood pressure or measuring your blood sugar yourself: frequency, time spent and inconvenience of this surveillance)	0.67 [0.52-0.79]	0.68 [0.51-0.79]	0.60 [0-0.91]
2C. Doctors visits (frequency and time spent for the visits)	0.66 [0.54-0.75]	0.65 [0.51-0.75]	0.74 [0.41-0.90]
2D. Arrange appointments and schedule doctors visits and lab tests	0.67 [0.57-0.75]	0.67 [0.54-0.77]	0.67 [0.41-0.87]
3. How would you rate the burden associated with taking care of paperwork from health insurance agencies, welfare organizations, hospitals and/or social care?	0.62 [0.50-0.73]	0.63 [0.48-0.74]	0.59 [0.17-0.83]
4. How would you rate the constraints associated with your diet (e.g., not being allowed to eat certain food)?	0.46 [0.30-0.60]	0.43 [0.27-0.60]	0.60 [0.14-0.89]
5. How would you rate the burden associated with the recommendations from your doctors to practice regular physical exercises?	0.48 [0.34-0.61]	0.48 [0.31-0.60]	0.52 [0.09-0.79]
6. What is the impact of your healthcare on your social relationships (e.g., need for assistance, being ashamed to take your medication in front of people)?	0.59 [0.44-0.70]	0.60 [0.43-0.73]	0.55 [0.13-0.82]
7. "Frequent healthcare reminds me of my health problems"	0.63 [0.52-0.73]	0.65 [0.51-0.74]	0.57 [0.19-0.82]
Global score*	0.76 [0.67-0.83]	0.75 [0.65-0.83]	0.78 [0.46-0.91]

Appendix 10(a): Reliability using test-retest (n=211) with ICC for agreement. 95% CIs were calculated by a bootstrap method. *Global score is the sum of all items scores with “Does not apply” and missing answers considered as the lowest possible score (0).

Appendix 10(b). Bland and Altman plot representing the test–retest reliability of the Treatment Burden Questionnaire global score (n=211)



4. Adapter la mesure du fardeau du traitement en anglais

Dans notre premier travail, nous avons développé un instrument de mesure du fardeau du traitement chez des patients souffrant de plusieurs maladies chroniques. Les items de l'instrument ont été développés puis validés auprès de patients ambulatoires ou hospitalisés, en France. Afin d'utiliser notre instrument dans un contexte ou une langue différente de ceux dans lesquels il a été développé, il est nécessaire de l'adapter. Cependant, l'adaptation linguistique et culturelle d'un instrument dépasse la simple traduction littérale. Des différences d'interprétation des items ou le besoin de questions différentes (i.e. items manquants ou inadaptés au nouveau contexte) peuvent, par exemple, rendre la mesure invalide si l'instrument n'est pas adapté correctement [74].

Dans ce chapitre, nous reviendrons d'abord sur les principes et les principales recommandations de l'adaptation linguistique et culturelle des mesures provenant du patient. Puis, nous présenterons notre deuxième travail portant sur l'adaptation et la validation de notre mesure du fardeau du traitement pour des patients vivant aux Etats-Unis, Royaume-Uni, Canada, Australie et Nouvelle-Zélande.

4.1. Adaptation linguistique et transculturelle

L'objectif de l'adaptation transculturelle est d'obtenir une « équivalence » entre les différentes versions d'un instrument. L'équivalence est un concept complexe, dont plus de 19 définitions différentes peuvent être retrouvées dans la littérature [75, 76]. Nous

présentons ici les définitions les plus fréquemment utilisées dans la littérature, d'après Herdman et al [76].

« *L'équivalence sémantique* » ou linguistique correspond à l'équivalence de formulation et de sens des items. Elle est en général obtenue et vérifiée par des méthodes de traduction. Cette définition est proche, pour la plupart des auteurs, de « l'équivalence fonctionnelle » qui représente la capacité de l'instrument adapté, dans la langue cible, à produire le même effet que l'instrument d'origine, dans la langue d'origine.

« *L'équivalence technique* » correspond au fait que l'administration de l'instrument est comparable entre les deux contextes. Cela comprend à la fois le mode d'administration (auto ou hétéro évaluation), la faisabilité et l'acceptabilité des items entre les différentes versions.

« *L'équivalence métrique* » et « *d'échelle* » correspondent au fait que les deux mesures - originale et adaptée- placent des individus similaires par rapport au concept mesuré sur le même point de l'échelle et aient des caractéristiques psychométriques comparables.

Afin d'assurer la meilleure équivalence possible entre échelles adaptées et originales, de nombreuses recommandations ont été établies [77-80]. Une revue systématique de la littérature synthétisant 17 recommandations différentes, a montré une grande homogénéité des méthodes recommandées pour l'adaptation d'une échelle dans une autre langue et culture [78]. Ce processus s'articule généralement autour de deux étapes : 1) adaptation du questionnaire dans la langue cible ; puis 2) évaluation des propriétés psychométriques de la version obtenue.

4.2. Recommandations pour l'adaptation transculturelle des mesures en santé

L'adaptation d'un instrument de mesure dans une langue différente de celle dans laquelle il a été développé comprend généralement 5 étapes : 1) traduction ; 2) concertation ; 3) rétro-traduction ; 4) harmonisation ; et 5) test de l'outil pré-final traduit.

La traduction consiste à traduire le questionnaire de la langue source vers la langue cible. La plupart des auteurs recommandent deux traducteurs au minimum. Ils doivent être bilingues et parler la langue cible couramment [80]. Idéalement, l'un des traducteurs doit connaître les concepts mesurés et l'autre non, ceci afin d'éviter l'utilisation de termes ambigus tout en assurant une équivalence conceptuelle[78].

La concertation consiste à synthétiser les différentes traductions obtenues précédemment. Cette synthèse est effectuée par un comité comprenant au moins les deux traducteurs et un observateur [78]. Les éventuelles divergences sont résolues par consensus. A la fin de cette étape, une unique version est adoptée.

La rétro-traduction consiste à traduire la version adoptée lors de l'étape précédente vers la langue originale. Les traducteurs impliqués sont bilingues, natifs du pays source. Ils n'ont ni participé à la première traduction, ni connaissance du questionnaire source.

L'harmonisation consiste à examiner l'ensemble des traductions obtenues dans les étapes précédentes. Un comité comprenant des méthodologistes, traducteurs et

professionnels de santé examinent les différences entre les versions et développent une version pré-finale.

L'outil pré-final obtenu est testé auprès un petit échantillon de patients. L'objectif est d'identifier des problèmes de compréhension, la répartition des réponses et la proportion de données manquantes associée à chaque item. Si cette étape produit des résultats satisfaisants alors on peut considérer que l'instrument a été correctement adapté dans la langue cible.

Une fois l'adaptation achevée, il est nécessaire de vérifier les propriétés métriques (i.e. validité et fiabilité) de l'instrument adapté. Ce processus doit être aussi rigoureux que possible, en suivant les mêmes recommandations que pour le développement d'un nouveau questionnaire. En effet, les propriétés métriques du questionnaire traduit ne sont pas automatiquement superposables à celles du questionnaire source. De subtiles différences peuvent avoir ainsi échappé aux différentes étapes d'adaptation et biaiser la mesure. Par exemple, l'adaptation en gallois du questionnaire de Beck a été réalisée en suivant toutes les étapes précédemment exposées et a abouti à une version « équivalente linguistiquement ». Cependant, lors de l'évaluation des propriétés psychométriques, les auteurs ont montré que les participants avaient un score de dépression significativement plus élevé en utilisant la version en gallois que lorsqu'ils utilisaient la version anglaise [81].

4.3. Adaptation et validation de notre outil de mesure du fardeau du traitement en anglais

L'objectif de notre deuxième travail était d'adapter l'outil de mesure du fardeau du traitement développé en France (TBQ), en anglais, afin de l'utiliser aux Etats-Unis, Royaume-Uni, Canada, Australie et Nouvelle-Zélande.

4.3.1. Méthodes

Nous avons adapté et validé le TBQ en anglais en suivant les recommandations précédemment exposées [77, 78, 80].

Après avoir adapté l'instrument en anglais en utilisant une méthode de traduction-rétro traduction « classique », nous avons vérifié la clarté et la faisabilité du questionnaire ainsi adapté. Pour cela, nous avons recruté 200 patients souffrant de maladies chroniques, tirés aléatoirement parmi les membres du site PatientsLikeMe. PatientsLikeMe est une communauté virtuelle regroupant plus de 200,000 patients volontaires souffrant de maladies chroniques. Les patients se connectent sur la plateforme, échangent entre eux et remplissent des questionnaires portant sur leurs maladies et leurs traitements. L'implication de patients lors du processus d'adaptation a permis de modifier le questionnaire en changeant des termes trop compliqués et en ajoutant deux nouveaux items absents dans la version française du TBQ : 1) le premier portant sur les contraintes financières du fardeau du traitement ; et 2) le second portant sur les difficultés que pouvaient avoir les patients avec les soignants (par exemple l'impression de ne pas être écoutés ou pris au sérieux par leurs soignants).

Dans un deuxième temps, nous avons validé l'outil adapté auprès d'un nouvel échantillon de patients souffrant d'au moins une maladie chronique, tirés

aléatoirement parmi les membres du site PatientsLikeMe. Nous avons ainsi exploré 1) la validité factorielle de notre instrument par l'étude du diagramme des valeurs propres d'une analyse factorielle ; 2) la validité de construit de l'instrument, en étudiant les associations entre le fardeau du traitement, l'observance thérapeutique mesurée à l'aide de l'échelle « Morisky Medication Adherence Scale » (MMAS-8) [82, 83], la qualité de vie mesurée à l'aide de l'échelle « PatientsLikeMe Quality of Life » (PLMQOL) [84], et la connaissance des patients vis-à-vis de leurs maladies et traitements ; et 3) la fiabilité de la mesure.

Les analyses statistiques ont été effectuées en utilisant SAS v9.2 (SAS Institute, Cary, NC, USA) et R v2.13.1 (<http://www.r-project.org>).

4.3.2. Résultats

Au total, 3000 patients ont été invités à participer à l'étude de validation et 610 (20.3%) patients ont répondu. L'âge moyen des participants était 51.5 (± 2.4) ans. La plupart des patients résidaient aux Etats-Unis d'Amérique (57.5%), au Royaume-Uni (8.7%), au Canada (8.4%) et en Australie/Nouvelle Zélande (3.4%).

La validité factorielle de notre instrument a été établie en étudiant le diagramme des valeurs propres d'une analyse factorielle. De même que dans l'outil original, nous avons retrouvé une structure unidimensionnelle du questionnaire.

Notre instrument a montré une bonne validité de construit ; nous avons mis en évidence : 1) une association significative entre fardeau du traitement et observance thérapeutique mesurée à l'aide du MMAS-8 (TBQ score=37.7 (SD=27.5) pour les patients avec une bonne observance thérapeutique vs. 61.8 (SD=30.5) pour les patients avec une mauvaise observance thérapeutique ($p < 0.0001$)) ; 2) une corrélation négative

entre fardeau du traitement et qualité de vie ($r_s=-0.5$, $p<0.0001$) et 3) une association significative entre fardeau du traitement et connaissance de la maladie (TBQ score=62.3 (SD=31.3) pour les patients jugeant avoir une connaissance insuffisante de leur maladie vs. TBQ score=47.8 (SD=30.4) pour les patients jugeant avoir une bonne connaissance de leur maladie ($p<0.0001$); et 4) une association significative entre fardeau du traitement et connaissance des traitements (TBQ score=63.0 (SD=31.6) pour les patients jugeant avoir une connaissance insuffisante de leurs traitements vs. TBQ score=49.3 (SD=30.7) pour les patients jugeant avoir une bonne connaissance de leurs traitements ($p<0.0001$)).

La reproductibilité de la mesure, évaluée à l'aide d'une méthode de test-retest, était bonne (ICC=0.77 [0.70-0.82], n=282).

Nous avons comparé les mesures du TBQ en fonction du pays de résidence des participants. De manière générale, les mesures étaient comparables entre les différents pays, à l'exception du fardeau financier du traitement qui était plus élevé aux Etats-Unis et en Australie/Nouvelle Zélande et des difficultés que pouvaient rencontrer les patients avec leurs soignants, qui était en moyenne plus élevées pour les patients résidant au Royaume-Uni.

4.3.3. Discussion

Nous avons adapté le TBQ en Anglais et testé les qualités métriques de l'instrument traduit, dans plusieurs pays anglophones. La mesure obtenue était valide et fiable dans les différents contextes et pays où elle a été testée.

Sur le plan thématique, cette étude nous a permis de mettre en valeur l'association entre le fardeau du traitement et : 1) l'inobservance thérapeutique; 2) une moindre

qualité de vie ; et 3) des connaissances insuffisantes sur leurs maladies et traitements, selon les patients. Ces associations avaient en effet été décrites sur le plan théorique mais jamais démontrées. Ces résultats soulignent la nécessité pour les médecins de s'intéresser au fardeau du traitement de leurs patients, afin d'améliorer leur observance thérapeutique et leur qualité de vie.

Sur le plan méthodologique, nous avons saisi l'opportunité de cette étude pour évaluer la faisabilité de l'adaptation d'un instrument de mesure dans une autre langue à l'aide d'une plateforme en ligne. Ceci nous a permis d'impliquer un grand nombre de patients (environ 800 en comptant la phase pilote), dans des pays distants géographiquement, en peu de temps (2 mois entre le début et la fin du recrutement). En revanche, il est probable que la population d'étude n'était pas représentative de la population des patients souffrant de maladies chroniques dans ces pays. D'autres études utilisant la plateforme PatientsLikeMe ont montré que les patients étaient des femmes, plus jeunes, et d'un plus haut niveau d'éducation [85-87]. Enfin, l'instrument adapté a été validé pour une administration en ligne, sur le site internet, même s'il est peu probable qu'un changement de format influe beaucoup sur les résultats.

Lors de l'adaptation du TBQ en anglais, nous avons ajouté deux items absents de la version d'origine. Le premier portait sur les contraintes financières du fardeau du traitement. En France, la sécurité sociale prend en charge les soins des patients souffrant de maladies chroniques et donc les patients évoquaient rarement ce problème. En revanche, dans les pays cibles, cela n'était pas toujours le cas. Par exemple, aux Etats-Unis, entre 30% et 47% des patients rapportent avoir des difficultés pour payer leurs soins [88]. Le second item que nous avons ajouté portait sur les

difficultés que pouvaient avoir les patients avec les soignants. Cet item n'avait pas été inclus dans l'outil original car nous pensions qu'il évaluait un concept différent. Cependant, il avait été mentionné dans plusieurs études qualitatives publiées depuis [47, 70, 71], et, dans la présente étude, les participants avaient insisté sur l'importance des discussions avec leurs médecins sur leurs connaissances vis-à-vis de leurs maladies et traitements et leur impact sur l'acceptabilité et l'observance des traitements. Il est possible que, lors de la première étude, les patients n'aient pas autant insisté sur ce point lors des interviews avec les investigateurs par biais de désirabilité sociale. Ainsi l'utilisation de la plateforme en ligne aurait permis aux patients de mieux exprimer leurs opinions vis-à-vis de sujets sensibles.

Le fardeau du traitement est un phénomène relativement homogène, quel que soit le pays de résidence des participants (**Table 4**). Cependant, l'addition de deux nouveaux items par rapport à la version française illustre le fait que des particularités culturelles, géographiques, ou liées aux systèmes de santé, influent sur le fardeau du traitement. Cette étroite relation entre fardeau du traitement et contexte de soin rend l'adaptation de notre outil dans d'autres langues et cultures complexe. Dans chaque contexte de soin, les items devraient être repensés, certains pourraient être ajoutés, d'autres modifiés ou supprimés. Ce processus impliquerait de multiples études dans de multiples pays. Et si certains pays partagent la même langue, ils ne partagent pas toujours la même culture ou le même système de soin. Ainsi, adapter notre instrument pourrait donc aboutir à un investissement considérable de temps, ressources et énergie. Par ailleurs, en ajoutant/retirant ou modifiant les items à chaque fois, comment la mesure resterait-elle comparable entre les différentes versions?

TBQ items	USA (n=351)	UK (n=53)	Canada (n=51)	Australia (n=21)	Other* (n=53)**
The taste, shape or size of your tablets and/or the annoyances caused by your injections (e.g., pain, bleeding, bruising or scars)	2,6 (3.0)	3,5 (2.9)	1,9 (2.7)	3 (3.1)	2,1 (2.9)
The number of times you should take your medication daily	2,4 (2.5)	3.0 (2.8)	2,7 (2.9)	1,4 (1.7)	2,2 (2.6)
The efforts you make not to forget to take your medications (e.g., managing your treatment when you are away from home, preparing and using pillboxes)	2,9 (2.9)	4.0 (3.2)	2,7 (2.6)	2,6 (2.7)	2,4 (2.9)
The necessary precautions when taking your medication (e.g., taking them at specific times of the day or meals, not being able to do certain things after taking medications such as driving or lying down)	3,2 (3.1)	4,2 (3.1)	3.0 (2.9)	2,9 (3.1)	2,7 (3.2)
Lab tests and other exams (e.g., blood tests or radiology): frequency, time spent and associated nuisances or inconveniences	3,1 (3.0)	3.0 (2.9)	1,9 (2.5)	3,2 (3.3)	2,7 (2.8)
Self-monitoring (e.g., taking your blood pressure or checking your blood sugar): frequency, time spent and associated nuisances or inconveniences	1,4 (2.5)	1.0 (2.1)	0,6 (1.1)	1,7 (2.8)	1,6 (2.5)
Doctor visits and other appointments: frequency and time spent for these visits and difficulties finding healthcare providers	3,9 (3.3)	3,8 (3.2)	2,9 (3.1)	4,1 (3.4)	3,3 (3.2)
The difficulties you could have in your relationships with healthcare providers (e.g., feeling not listened to enough or not taken seriously)	3,7 (3.5)	5,4 (3.8)	3,6 (3.5)	3,6 (3.5)	3,5 (3.5)
Arranging medical appointments and/or transportation (doctors visits, lab tests and other exams) and reorganizing your schedule around these appointments	3,8 (3.4)	4,4 (3.4)	2,6 (2.6)	4,2 (3.5)	2,9 (3.0)
The administrative burden related to healthcare (e.g., all you have to do for hospitalizations, insurance reimbursements and/or obtaining social services)	3,8 (3.4)	2,8 (3.4)	2,1 (2.5)	3,8 (3.4)	2,8 (3.1)
The financial burden associated with your healthcare (e.g., out-of-pocket expenses or expenses not covered by insurance)	5,6 (3.6)	4.0 (3.9)	3,8 (3.6)	5,1 (3.7)	4 (3.7)
The burden related to dietary changes (e.g., avoiding certain foods or alcohol, having to quit smoking)	3,1 (3.3)	4.0 (3.6)	3,3 (3.4)	3,8 (4.0)	3,6 (3.6)
The burden related to doctors' recommendations to practice physical activity (e.g., walking, jogging, swimming)	3,7 (3.4)	4,6 (3.8)	3,1 (3.2)	4,3 (3.7)	3,2 (3.6)
How does your healthcare impact your relationships with others (e.g., being dependent on others and feeling like a burden to them, being embarrassed to take your medications in public)	4,9 (3.7)	5,7 (3.7)	4 (3.5)	6,7 (3.8)	4,1 (3.4)
' The need for medical healthcare on a regular basis reminds me of my health problems '	6,1 (3.6)	5,2 (3.8)	5,3 (3.7)	6,4 (3.4)	5,4 (3.5)

Table 6 : Comparaison des scores du fardeau du traitement mesurés par le TBQ en fonction du pays de résidence des participants

4.3.4. Article

Tran VT, Harrington M, Montori VM, Barnes C, Wicks P, Ravaud P. Adaptation and validation of the Treatment Burden Questionnaire (TBQ) in English using an internet platform. *BMC Med.* 2014;12:109.

RESEARCH ARTICLE

Open Access

Adaptation and validation of the Treatment Burden Questionnaire (TBQ) in English using an internet platform

Viet-Thi Tran^{1,2,3*}, Magdalena Harrington⁴, Victor M Montori⁵, Caroline Barnes^{2,3,6}, Paul Wicks⁴ and Philippe Ravaud^{2,3,6,7}

Abstract

Background: Treatment burden refers to the workload imposed by healthcare on patients, and the effect this has on quality of life. The Treatment Burden Questionnaire (TBQ) aims to assess treatment burden in different condition and treatment contexts. Here, we aimed to evaluate the validity and reliability of an English version of the TBQ, a scale that was originally developed in French.

Methods: The TBQ was translated into English by a forward-backward translation method. Wording and possible missing items were assessed during a pretest involving 200 patients with chronic conditions. Measurement properties of the instrument were assessed online with a patient network, using the PatientsLikeMe website. Dimensional structure of the questionnaire was assessed by factor analysis. Construct validity was assessed by associating TBQ global score with clinical variables, adherence to medication assessed by Morisky's Medication Adherence Scale (MMAS-8), quality of life (QOL) assessed by the PatientsLikeMe Quality of Life Scale (PLMQOL), and patients' confidence in their knowledge of their conditions and treatments. Reliability was determined by a test-retest method.

Results: In total, 610 patients with chronic conditions, mainly from the USA, UK, Canada, Australia, or New Zealand, completed the TBQ between September and October 2013. The English TBQ showed a unidimensional structure with Cronbach α of 0.90. The TBQ global score was negatively correlated with the PLMQOL score ($r_s = -0.50$; $p < 0.0001$). Low rather than moderate or high adherence to medication was associated with high TBQ score (mean [SD] TBQ score 61.8 [30.5] vs. 37.7 [27.5]; $P < 0.0001$). The treatment burden was higher for patients who had insufficient knowledge compared with those who had sufficient knowledge about their treatments (mean \pm SD TBQ score 62.3 \pm 31.3 vs. 47.8 \pm 30.4; $P < 0.0001$) and conditions (63.0 \pm 31.6 vs. 49.3 \pm 30.7; $P < 0.0001$). The intraclass correlation coefficient for the retest ($n = 282$) was 0.77 (95% CI 0.70 to 0.82).

Conclusions: We found that the English TBQ is a reliable instrument in this population, and provide evidence supporting the construct validity for its use to assess treatment burden for patients with one or more chronic conditions in English-speaking countries.

Keywords: Chronic Disease/therapy, Patient Participation, Quality of Life, Medication Adherence, Questionnaires

* Correspondence: thi.tran-viet@htd.aphp.fr

¹Department of General Medicine, Paris Diderot University, Paris, France

²METHODS Team, Epidemiology and Statistics Sorbonne Paris Cité Research Centre, UMR 1153, INSERM, 1 place du Parvis Notre-Dame, Paris 75181, France

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



© 2014 Tran et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. 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.

Background

Treatment burden is defined as the 'work' of being a patient and its effect on the quality of life (QOL) of patients. It represents the challenges associated with everything patients have to do to take care of themselves (e.g. medication intake, drug management, self-monitoring, visits to the physician, laboratory tests, lifestyle changes, administrative tasks to access and coordinate care) [1-3]. This work can represent a tremendous investment of time, attention, cognitive energy, and effort. For example, a patient with a chronic condition may spend 2 hours or more per day on health-related activities [4], and this increases for patients with multiple chronic conditions who receive treatment or recommendations for each of their diseases [5,6]. For these patients, treatment burden could be considered a crucial outcome for disease management [7]. Difficulties with their treatment are often not shared in depth by patients during medical consultations [8], and physicians are often not aware of the challenges their patients face in coping with everything asked of them [9]. Thus, a tool that is simple to understand and easy to administer is needed to identify overburdened patients in daily practice and for research, in order to design new therapeutic strategies that are both efficient and acceptable for patients.

The Treatment Burden Questionnaire (TBQ) is the only existing instrument that measures treatment burden without restricting its scope to a single condition or treatment context [9,10]. The TBQ is composed of 13 items rated on a Likert scale ranging from 0 (not a problem), to 10 (big problem). The TBQ was derived from a literature review and qualitative semi-structured interviews with patients in France. It assesses the burden associated with taking medicine, self-monitoring, laboratory tests, doctor visits, need for organization, administrative tasks, following advice on diet and physical activity, and social impact of the treatment. Item scores can be summed into a global score, ranging from 0 to 130. The instrument was first developed with a sample of patients with one or more chronic conditions, who were recruited from hospitals and general-practice clinics in France.

To measure treatment burden in English-speaking countries, the TBQ needs to be adapted: items must remain comprehensible without changing their original meaning, and it should be recognized that new items could arise in different contexts whereas some existing items might be irrelevant. Patients must be involved in the process to ensure that the resulting tool is clear and relevant [11].

In this study, we aimed to assess the validity and reliability of an English version of the TBQ intended to be used in both routine practice and in research for patients with at least one chronic condition living in the USA, Canada, UK, Australia, or New Zealand.

Methods

Adaptation and validation of the TBQ in English followed a multi-step approach recommended in the literature [12-14]: 1) review the conceptual evidence about treatment burden in English-speaking countries; 2) translate the TBQ into English; 3) pretest the instrument with patients to assess the relevance of items, clarity and wording and; 4) assess validity and reliability by a test-retest of the adapted instrument.

Conceptual evidence of treatment burden in English-speaking countries

During a literature review of MEDLINE via Pubmed, we identified several articles describing the concept of treatment burden in English-speaking countries. In the USA, Eton and colleagues identified three broad themes for treatment burden: 1) work patients must do to take care of their health, 2) problem-focused strategies to facilitate self-care, and 3) factors that exacerbate the perceived burden [2]. In the UK, Gallacher and associates found four similar themes in patients with chronic heart failure and stroke: 1) learning about treatments and their consequences, 2) engaging with others, 3) adhering to treatment and lifestyle changes, and 4) monitoring their treatments [15,16]. Finally, in Australia, Sav and associates identified four inter-related components of treatment burden: 1) financial, 2) time and travel, 3) medication, and 4) healthcare access burdens [17].

The original TBQ encompassed all of these domains, with the exception of the financial treatment burden, because in France, the public health insurance program guarantees healthcare free of charge for patients with chronic conditions. Therefore, we added a new item in the English adaptation of the TBQ: 'How would you rate the financial burden associated with your healthcare (e.g., out-of-pocket expenses or expenses not covered by insurance)?'

Translation of the TBQ into English

Translation of the TBQ into English involved a classic 'forward-backward' translation method [13]. First, the original instrument in French was translated into English by two bilingual translators; one (CB) had a medical background, and was familiar with the concept of treatment burden. Second, the two obtained translations were synthesized and reviewed by a committee, which included the authors of the original questionnaire. Third, two different translators, blinded to the original version, back-translated the questionnaire into French. Finally, the committee reviewed and synthesized all translations to elaborate English items that were similar to the original items and easy to answer.

Pre-testing the instrument

To assess the relevance of items, clarity, and wording, we pre-tested the obtained instrument with a convenience

sample of 200 participants in August 2013 (see Additional file 1). We used an internet platform, the Open Research Exchange (ORE) [18,19], to recruit patients on PatientsLikeMe (PLM) [20], an online network where 200,000 voluntary participants with chronic conditions share data about their treatment, conditions, and symptoms. Members of PLM join the site with the expectation that they will be participating in research. To participate, patients had to have at least one chronic condition (defined as requiring ongoing healthcare for at least 6 months). After having answered the questionnaire, they provided feedback about 1) the clarity, wording, and relevance of the items, and 2) any burden that they felt was not covered or insufficiently covered in the questionnaire in an open-ended manner. Their answers were categorized and discussed by two of the authors (VT-T and CB).

Concerning the wording of items, 10 patients (5.0%) felt that the word 'constraints' was confusing, thus we replaced the word 'constraints' by the word 'problem', as suggested by the patients. Patients were also asked whether there were any important elements of treatment burden that they considered to be missing from questions: 15 patients (7.5%) thought that relationships between patients and healthcare providers were insufficiently covered in the original items. Other suggestions were either specific to a particular condition, were related to the burden of disease, or were already covered in the existing items. Thus, we added a new item for testing: 'How would you rate the difficulties you could have in your relationships with healthcare providers (e.g., feeling not listened to enough or not taken seriously)?' After the pretest, the English TBQ was therefore composed of 15 items, with rating scales ranging from 0 to 10 and labeled anchors ('not a problem' and 'large problem').

Assessment of validity and reliability of the English TBQ

We studied the measurement properties of the instrument by 1) describing the item properties, 2) assessing factor structure, 3) assessing construct validity, and 4) assessing reliability by test-retest.

We recruited a convenience sample of patients via the aforementioned internet platform. Patients were eligible if they were 18 years or older, and had at least one condition that had required ongoing health care for at least 6 months. We sent email invitations to a random sample of 3,000 members on the internet platform who did not participate in the pretest and who met the eligibility criteria, encouraging them to connect to the website and complete the questionnaire. To increase the number of respondents, an email reminder was sent after 2 months. Patients consented electronically to participate in the study. The recruitment message outlined the purpose of the study and reminded patients that they were under no obligation to participate, that their aggregated results

may be published. Because there were no anticipated adverse consequences for participation, institutional review board (IRB) approval was not sought for this project.

Item properties were described using three criteria: 1) proportion of missing answers, 2) relevance of items assessed by the proportion of 'does not apply', and 3) score distributions.

Factor structure was investigated by exploratory factorial analysis. Scree plots were used to visualize a break between factors with large eigenvalues and those with smaller eigenvalues. Factors that appeared before the horizontal break were assumed to be meaningful. Internal consistency was assessed by Cronbach's α [21], and was considered acceptable between 0.70 and 0.95 [22].

The global score of the TBQ (TBQ Global score) was the sum of the answers to each item. 'Does not apply' or missing answers were considered the lowest possible score (0) because we considered that a patient not concerned by a domain of treatment burden had no burden for that domain.

Construct validity was tested by confirming four pre-specified hypotheses. First, we expected a negative correlation between treatment burden (as measured by the TBQ global score) and quality of life. Quality of life was measured by the PatientsLikeMe Quality of Life (PLMQOL) scale, a validated 24-item questionnaire assessing physical, mental, and social quality of life. PLMQOL scores range from 0 to 100 for each domain (higher scores indicating better quality of life) and are summed for a global assessment of quality of life [23]. Second, we predicted an association between TBQ global score and adherence to medication: the greater the treatment burden, the lower the adherence to treatment. Adherence to medical treatment was measured by Morisky's Medication Adherence Scale 8 (MMAS-8) [24,25], a validated eight-item questionnaire, with scores ranging from 0 to 8. High adherence is a score of 8; medium adherence, 6 to 7; and low adherence, less than 6 [24]. Third, we hypothesized that patients with better knowledge of their conditions and treatments would have a low treatment burden. Confidence in patients' knowledge about their treatments and conditions was assessed by two questions: 'Do you think you have sufficient knowledge about your conditions (e.g., symptoms, disease progression)?' and 'Do you think you have sufficient knowledge about your treatments (e.g., possible side effects, expected benefits, other treatment options)?'. Answers were rated on a five-step scale: 'very sufficient', 'sufficient', 'average', 'insufficient' and 'very insufficient'. Finally, we assumed a positive correlation between TBQ global score and the following clinical variables: 1) number of conditions, 2) drug administration (number of tablets, injections, and administrations per day), and 3) medical follow-up (number of different physicians, medical appointments per month, and hospitalizations per year).

To elicit the chronic conditions a patient had, we asked the patient to self identify the condition(s) from a list recommended as core for any measure of multimorbidity [26]. Options were presented as categories illustrated by common conditions; for example: 'Rheumatologic disease (e.g. osteoporosis, arthritis, or inflammatory polyarthropathies)'. Patients were encouraged to complete their answer with free text. The text was analyzed, and the condition was categorized in the appropriate category by a single investigator (VTT).

The association between the TBQ global score, quality of life score and clinical variables was assessed by Spearman correlation coefficient (r_s), which was considered high when greater 0.50, and moderate when 0.35 to 0.50 [27]. Wilcoxon and Kruskal-Wallis tests were used to compare qualitative variables across groups. $P < 0.05$ was considered statistically significant.

Reliability of the instrument was determined by a test-retest method. Patients were asked to complete the new instrument twice: at baseline and again after 2 weeks when they received an email reminder. Reliability was assessed by the intraclass correlation coefficient (ICC) for agreement [28], defined as the ratio of the subject variance by the sum of the subject variance, the rater variance and the residual. The 95% confidence interval (CI) was determined by a bootstrap method. Agreement was considered acceptable when ICC was greater than 0.60 [29]. Agreement was represented by Bland and Altman plots, which represent the differences between two measurements against the means of the two measurements [30].

Statistical analyses were performed using SAS (version 9.3; SAS Inst., Cary, NC, USA) and R (version 3.0 [31], the R Foundation for Statistical Computing, Vienna, Austria).

Results

In total, 3,000 patients were invited to complete the TBQ, and 610 (20.3%) did so between September and October 2013. The mean \pm SD age was 51.5 ± 2.4 years. Patients resided mainly in the USA (57.5%), Canada (8.4%), UK (8.7%), and Australia/New Zealand (3.4%). Some patients (13.3%) lived in other countries (Belgium, Brazil, Chile, Czech Republic, Denmark, Finland, France, Germany, India, Ireland, Israel, Italy, Malaysia, Norway, Oman, Portugal, Russia, South Africa, Spain, Switzerland, Ukraine, and Uzbekistan), or the country of residence was missing (8.7%). The mean \pm SD number of chronic conditions was 2.9 ± 1.9 (range 1 to 11) (Table 1).

We found a high completion rate (>99%) for all items. All items were relevant for all patients, with the proportion of 'does not apply' ranging from 0.6% to 7.2%, with the exception of the burden associated with self-monitoring (30.3%) (see Additional file 2). The TBQ global score was the sum of the answers to each of the 15 items, and ranged from 0 to 150 ('Does not apply' was considered the lowest

possible score). The global score had a high correlation with the score for each item of the TBQ ($r_s = 0.47$ to 0.71) (see Additional file 3).

Factor structure, assessed by scree plots, favored a unidimensional instrument. The first factor explained 87% of the variance, and had an eigenvalue of 5.91. For subsequent factors, eigenvalues were less than 0.80 (see Additional file 4).

Construct validity showed a significant moderate negative correlation between the TBQ global score and PLMQOL score ($r_s = -0.50$; $P < 0.0001$). Correlation coefficients ranged from $r_s = -0.39$ ($P < 0.0001$) for physical quality of life to $r_s = -0.50$ ($P < 0.0001$) for mental quality of life, indicating that patients with high TBQ score had low quality of life (Table 2). We found a significant association between treatment burden measured by the TBQ global score and adherence to medication measured by the MMAS-8: mean \pm SD TBQ global score was 37.7 ± 27.5 for patients with high or moderate adherence, and 61.8 ± 30.5 for patients with low adherence ($P < 0.0001$) (Table 2). The TBQ score was also significantly associated with patients' confidence in their knowledge about their treatments and conditions. The mean \pm SD TBQ global score was 47.8 ± 30.4 for patients who felt that they had sufficient knowledge about their treatment, versus 62.3 ± 31.3 for those who felt that they had insufficient knowledge ($P < 0.0001$). The mean \pm SD TBQ global score was 49.3 ± 30.7 for patients declaring sufficient knowledge about their conditions versus 63.0 ± 31.6 for those with insufficient knowledge ($P < 0.0001$) (Table 2). Finally, we found a low to moderate significant positive correlation of TBQ score with clinical variables in terms of number of different conditions, drug administrations (number of tablets, injections, and administrations per day) and medical follow-up (number of different physicians, medical appointments per month, and hospitalizations per year) (see Additional file 5).

Reliability was assessed by a test-retest method for 282 patients (46%). Participants in the retest showed similar characteristics as patients in the test phase (Table 1). For the TBQ global score, the ICC for agreement was 0.77 (95% CI 0.70 to 0.82) (see Additional file 6). Bland and Altman plots for the global TBQ score showed a mean difference of -0.58 ; 95% limits of agreement -43.5 and 42.3 (Figure 1).

Subgroup analysis showed a high TBQ global score for patients who needed equipment such as wheelchairs or canes (mean \pm SD score 69.0 ± 33.4), those who received physical therapy (62.7 ± 33.9), and those who had gastrointestinal (65.4 ± 32.5) or skin (64.9 ± 30.8) diseases (see Additional file 7). TBQ scores did not differ for the most common healthcare location (public hospital, private hospital, general practice, or specialist clinic) ($P = 0.97$).

Overall, treatment burden scores were homogeneous and comparable between countries. The lowest mean (SD)

Table 1 Demographic and clinical characteristics of patients in the validation study (n = 610) and retest of the Treatment Burden Questionnaire (TBQ) for an English-speaking population (n = 282)

Characteristic	Validation study (n = 610)	Retest (n = 282)
Age, years ^a	51.5 ± 12.4	52.3 ± 12.3
Female sex, n (%)	451 (76.7)	215 (76.2)
Country of residence, n (%)		
USA	351 (57.5)	172 (61.0)
Canada	51 (8.4)	20 (7.1)
UK	53 (8.7)	26 (9.2)
Australia/New Zealand	21 (3.4)	10 (3.5)
Other/missing	134 (22.0)	54 (19.1)
Educational level, n(%)		
Less than High School	24 (3.9)	10 (3.5)
High school graduate or General Education Diploma	99 (16.2)	55 (19.5)
Some College	220 (36.1)	95 (33.7)
Bachelor's degree	134 (22.0)	60 (21.3)
Graduate degree	124 (20.3)	58 (20.6)
Treatments, n		
Tablets and pills/day	8.5 ± 6.4	8.6 ± 6.7
Injections/week	1.4 ± 4.6	1.5 ± 4.1
Drug administrations/day	3.0 ± 2.0	2.9 ± 1.9
Different doctors the patient sees	3.0 ± 2.3	2.9 ± 1.5
Appointments/month	2.9 ± 2.9	2.7 ± 2.7
Hospitalizations/year	0.5 ± 1.7	0.5 ± 1.1
Presence of an informal caregiver, n (%)	280 (45.9)	125 (44.3)
Most common location for medical consultations, n (%)		
Public hospital	63 (10.3)	30 (10.6)
Private hospital	20 (3.3)	4 (1.4)
General practice clinic	291 (47.7)	135 (47.9)
Specialist clinic	224 (36.7)	106 (37.6)
Duration of oldest chronic condition, years, n (%)		
<5	182 (29.8)	75 (26.6)
5 to 10	217 (35.6)	108 (38.3)
> 10	205 (33.6)	96 (34.0)
Chronic conditions, n	2.9 ± 1.9	2.9 ± 1.9
Conditions, n (%) ^b		
Neurologic disease	277 (45.4)	134 (47.5)
Psychiatric disease	250 (41.0)	107 (37.9)
Rheumatologic disease	203 (33.3)	89 (31.6)
High blood pressure	156 (25.6)	62 (22.0)
Gastrointestinal disease	129 (21.1)	67 (23.7)
Endocrine disorder (other than diabetes)	121 (19.8)	53 (18.8)

Table 1 Demographic and clinical characteristics of patients in the validation study (n = 610) and retest of the Treatment Burden Questionnaire (TBQ) for an English-speaking population (n = 282) (Continued)

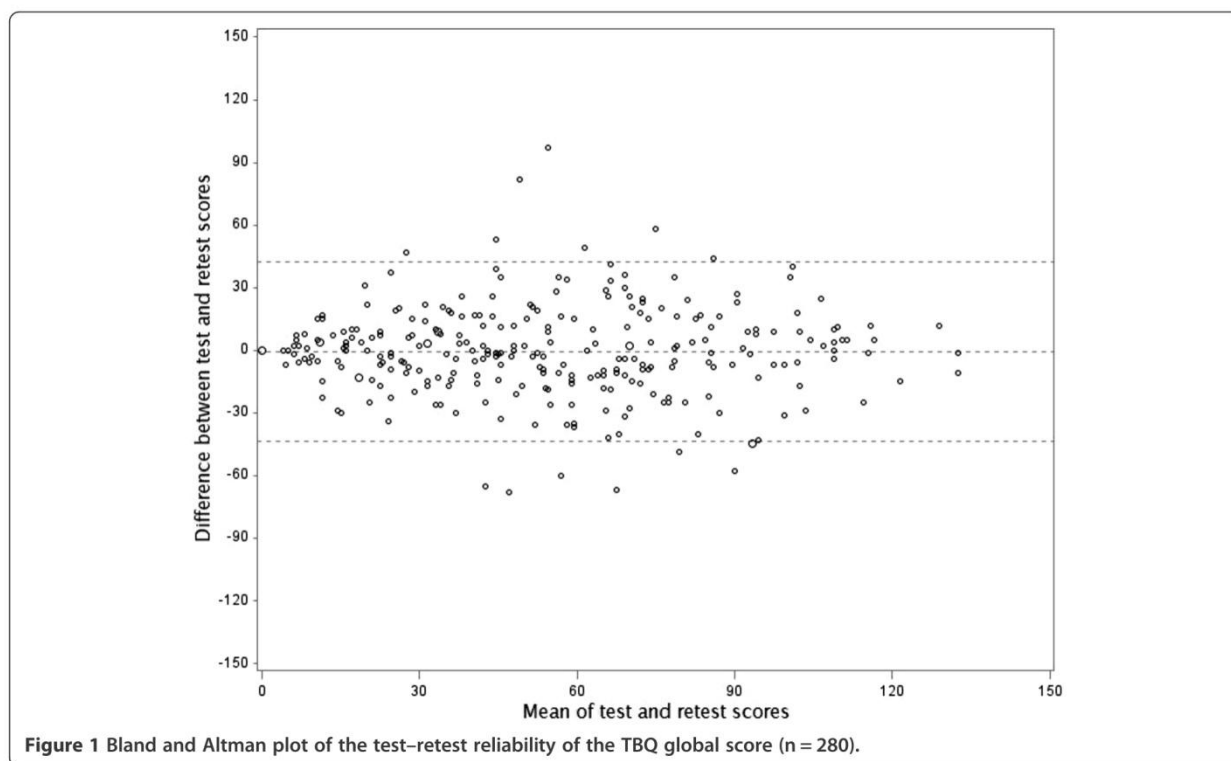
Lung disease	93 (15.2)	45 (16.0)
Vision problems	83 (13.6)	35 (12.4)
Fibromyalgia	79 (12.9)	30 (10.6)
Skin disease	71 (11.6)	33 (11.7)
Hearing problem	49 (8.0)	27 (9.6)
Diabetes	45 (7.4)	19 (6.7)
Kidney disease	38 (6.2)	12 (4.2)
Heart disease	34 (5.6)	14 (5.0)
Cancer or malignant blood disease	31 (5.1)	19 (6.7)
Infectious disease	19 (3.1)	10 (3.5)
Stroke or cerebrovascular disease	17 (2.8)	8 (2.8)

^aData are mean ± SD unless otherwise indicated.
^bAny given patient can have multiple conditions.

Table 2 Association of adherence to medication, patients' evaluation of their knowledge about their treatment and conditions and QOL with TBQ global score (n = 610)

	TBQ global score, mean ± SD	Spearman correlation coefficient	P value
Adherence to treatment assessed by the MMAS-8 ^a			
High and moderate adherence	37.7 ± 27.5	–	<0.0001
Low adherence	61.8 ± 30.5	–	<0.0001
Patients' evaluation of their knowledge of:			
Their treatments ^b			
Sufficient	47.8 ± 30.4	–	<0.0001
Insufficient	62.3 ± 31.3	–	<0.0001
Their conditions ^b			
Sufficient	49.3 ± 30.7	–	<0.0001
Insufficient	63.0 ± 31.6	–	<0.0001
Quality of life assessed by the PLMQOL ^c			
Physical QOL	–	–0.39	<0.0001
Mental QOL	–	–0.50	<0.0001
Social QOL	–	–0.49	<0.0001
Overall assessment of QOL	–	–0.50	<0.0001

PLMQOL, PatientsLikeMe Quality of Life; QOL, quality of life; TBQ, Treatment Burden Questionnaire.
^aHigh adherence is a score of 8; medium adherence, 6 to 7; and low adherence, <6 [24]. ^b'Very sufficient' and 'sufficient' was considered sufficient and 'average', 'insufficient' and 'very insufficient' was considered insufficient.
^cPLMQOL assesses physical, mental and social QOL; high scores indicate better QOL, while negative correlation coefficients indicate that patients with high treatment burden have low QOL.



TBQ score was for the item related to self-monitoring (from 0.6 ± 1.1 in Canada to 1.7 ± 2.8 in Australia/New Zealand) and the highest was for the item related to the effect of healthcare on relationships with others (from 4.0 ± 3.5 in Canada to 6.7 ± 3.8 for Australia/New Zealand) (Figure 2).

Discussion

We adapted the TBQ for patients with one or more chronic conditions living in the USA, Canada, the UK, Australia, or New Zealand. This is the first valid and reliable instrument assessing treatment burden across multiple conditions and treatments.

During the adaptation process, we added two new items to the English TBQ that did not exist in the original study, which took place in France [9]. First, we added an item about the financial treatment burden. In France, the national health program guarantees healthcare free of charge for patients with chronic conditions. However, in countries where the English TBQ was administered, this is not always the case [10,17]. For example, 30% to 47% of patients in the USA may have problems paying medical bills [32]. Second, during the pretest, patients expressed difficulties that they had in their relationships with healthcare providers. This item was not included in the original French questionnaire because the authors felt that it measured a concept different from treatment burden. However, it was mentioned in subsequent qualitative studies [2,15,16] and, in the present

study, participants insisted on including this item to take into account how discussion with physicians could affect how much they knew about their conditions and treatment, and how non-optimal relationships with healthcare providers could hinder adherence to treatment, lifestyle changes, or follow-up. During the first study, patients may have felt hindered in mentioning this issue to the investigators during face-to-face interviews and therefore the point was underestimated. Use of a web-based questionnaire has shown a decrease in socially desirable responses, and an increase in willingness to disclose sensitive information [33].

Our results highlight the advantages of using an internet platform: we were able to include 200 patients with chronic conditions from different countries in the pretest in less than 1 month. Therefore, we gathered rich qualitative data on questionnaire items and possible missing concepts, indicating the high content validity of our tool.

The item properties of the English TBQ were comparable with those from the original study in France with 1) a significant floor effect (>15%) for all items, which reflects that patients could have no burden in the aspects of their care that they have already integrated into their lives, and 2) about 30% of 'Does not apply' for the item related to self-monitoring. As for the original questionnaire in French, we chose to keep the item about self-monitoring because it was considered a mandatory item for other patients, particularly those with diabetes.

TBQ items	USA (n=351)	United Kingdom (n=53)	Canada (n=51)	Australia/New Zealand (n=21)	Other* (n=53)**
The taste, shape or size of your tablets and/or the annoyances caused by your injections (e.g., pain, bleeding, bruising or scars)	2,6 ± 3,0	3,5 ± 2,9	1,9 ± 2,7	3 ± 3,1	2,1 ± 2,9
The number of times you should take your medication daily	2,4 ± 2,5	3,0 ± 2,8	2,7 ± 2,9	1,4 ± 1,7	2,2 ± 2,6
The efforts you make not to forget to take your medications (e.g., managing your treatment when you are away from home, preparing and using pillboxes)	2,9 ± 2,9	4,0 ± 3,2	2,7 ± 2,6	2,6 ± 2,7	2,4 ± 2,9
The necessary precautions when taking your medication (e.g., taking them at specific times of the day or meals, not being able to do certain things after taking medications such as driving or lying down)	3,2 ± 3,1	4,2 ± 3,1	3,0 ± 2,9	2,9 ± 3,1	2,7 ± 3,2
Lab tests and other exams (e.g., blood tests or radiology): frequency, time spent and associated nuisances or inconveniences	3,1 ± 3,0	3,0 ± 2,9	1,9 ± 2,5	3,2 ± 3,3	2,7 ± 2,8
Self-monitoring (e.g., taking your blood pressure or checking your blood sugar): frequency, time spent and associated nuisances or inconveniences	1,4 ± 2,5	1,0 ± 2,1	0,6 ± 1,1	1,7 ± 2,8	1,6 ± 2,5
Doctor visits and other appointments: frequency and time spent for these visits and difficulties finding healthcare providers	3,9 ± 3,3	3,8 ± 3,2	2,9 ± 3,1	4,1 ± 3,4	3,3 ± 3,2
The difficulties you could have in your relationships with healthcare providers (e.g., feeling not listened to enough or not taken seriously)	3,7 ± 3,5	5,4 ± 3,8	3,6 ± 3,5	3,6 ± 3,5	3,5 ± 3,5
Arranging medical appointments and/or transportation (doctors visits, lab tests and other exams) and reorganizing your schedule around these appointments	3,8 ± 3,4	4,4 ± 3,4	2,6 ± 2,6	4,2 ± 3,5	2,9 ± 3,0
The administrative burden related to healthcare (e.g., all you have to do for hospitalizations, insurance reimbursements and/or obtaining social services)	3,8 ± 3,4	2,8 ± 3,4	2,1 ± 2,5	3,8 ± 3,4	2,8 ± 3,1
The financial burden associated with your healthcare (e.g., out-of-pocket expenses or expenses not covered by insurance)	5,6 ± 3,6	4,0 ± 3,9	3,8 ± 3,6	5,1 ± 3,7	4 ± 3,7
The burden related to dietary changes (e.g., avoiding certain foods or alcohol, having to quit smoking)	3,1 ± 3,3	4,0 ± 3,6	3,3 ± 3,4	3,8 ± 4,0	3,6 ± 3,6
The burden related to doctors' recommendations to practice physical activity (e.g., walking, jogging, swimming)	3,7 ± 3,4	4,6 ± 3,8	3,1 ± 3,2	4,3 ± 3,7	3,2 ± 3,6
How does your healthcare impact your relationships with others (e.g., being dependent on others and feeling like a burden to them, being embarrassed to take your medications in public)	4,9 ± 3,7	5,7 ± 3,7	4 ± 3,5	6,7 ± 3,8	4,1 ± 3,4
'The need for medical healthcare on a regular basis reminds me of my health problems'	6,1 ± 3,6	5,2 ± 3,8	5,3 ± 3,7	6,4 ± 3,4	5,4 ± 3,5

Item scores are presented as mean (SD). Highest scores are in red and lowest in green. *Other includes Belgium, Brazil, Chile, Czech Republic, Denmark, Finland, France, Germany, India, Ireland, Israel, Italy, Malaysia, Norway, Oman, Portugal, Russia, South Africa, Spain, Switzerland, Ukraine, Uzbekistan. **Participants with missing country of residence (n=81) were excluded from the analysis

Figure 2 Comparison of scores for TBQ items by country (n = 529)^a.

The validation study included 610 patients from different countries with different conditions and treatments. As hypothesized, we highlighted statistically significant associations between the overall treatment burden and adherence to medication, patients' confidence in their knowledge about their conditions and treatment, and quality of life. To our knowledge, this is the first study to report such associations independent of disease or treatment contexts. Development of interventions to improve patient knowledge about therapeutic options could lead to shared decision making, reduced treatment burden, better adherence to medication, and better quality of life for patients.

We compared TBQ scores between countries and found that most item scores were homogeneous for patients in the USA, UK, Canada, and Australia/New Zealand, with two exceptions. First, we found increased financial treatment burden for patients residing in the USA and Australia. This result was expected because qualitative studies in these countries described the financial burden as one of the issues most widely discussed by patients [10,17]. Second, in the UK, we found increased treatment burden related to patients' relationships with their healthcare providers. This finding could be related to pay-for-performance initiatives (for example, Quality and Outcomes Framework) that may encourage physicians to focus on biomedical objectives and spend less time listening to patients' non-medical problems. More research is

needed to understand the precise aspects of treatment burden, and why patients might experience different aspects in each of these countries.

This study has several limitations. First, to measure patients' confidence in their knowledge about their treatments and conditions, we used questions that were not validated. However, participants reviewed these questions during the pretest in the same way they reviewed the questions of the TBQ. Patients expressed no difficulties in answering these questions, and we found neither any missing answers, nor a floor or ceiling effect. Second, we recruited patients through the PLM website only, and from this patient group, only 20% of the invited patients responded to the web survey. This response rate was similar to that from other studies involving online surveys. This recruitment method might have selected patients willing to share their experiences with others and/or with severe or rare conditions. We did not include participants who were not able to access a computer. Therefore, our sample is not representative of the general population of patients with one or more chronic conditions: overall patients in our study were younger, more educated, and more often female (similar to other studies involving the PLM website [34-36]). Prevalence of chronic conditions in our study differed from those of chronic conditions of patients seeking healthcare in national surveys. For example, we found

a high proportion of patients with neurological conditions, because PLM was initially developed for these patients. Although these biases were considered relatively small, psychometric properties for the TBQ (especially reliability estimates) might differ in other populations. Despite these limitations, use of an online platform has a number of advantages. First, we were able to involve a larger number of patients than in usual cross-cultural adaptation studies (800 participants including pretest) [37] in a very short time (2 months from pretest to results). Second, the recruited patients were geographically dispersed and thus, we might have avoided selecting only patients living near research centers. Finally, patients involved were engaged in research, and therefore provided us with rich feedback at every stage of our research, which may have contributed to the development of questionnaire items closer to patients' own words and problems.

The present study advances the evidence of the validity of the TBQ. Further studies should focus on testing other psychometric properties such as responsiveness in longitudinal and prospective studies. The TBQ can be implemented in clinical trials assessing interventions intended to mitigate treatment burden and its negative effect on quality of life to develop a Minimally Disruptive Medicine [1].

Conclusions

We adapted and validated the TBQ for English-speaking countries (USA, UK, Canada, and Australia) using an online patient-based platform for direct interaction with patients. This study resulted in the first instrument assessing treatment burden for any condition and treatment context in these countries.

Additional files

Additional file 1: Demographic and clinical characteristics of patients included in the pretest of the Treatment Burden Questionnaire (TBQ) (n = 200).

Additional file 2: Item properties of the Treatment Burden Questionnaire (TBQ) (n = 610 patients).

Additional file 3: Correlation of each item of the Treatment Burden Questionnaire (TBQ) with the global score, omitting that item from the total (n = 610).

Additional file 4: Scree plots and eigenvalues for the correlation matrix of the factor analysis (n = 610).

Additional file 5: Correlation between treatment burden scores and clinical variables in terms of treatment workload (n = 610).

Additional file 6: Reliability testing by a test-retest method (n = 282).

Additional file 7: Treatment Burden Questionnaire (TBQ) global score by different subgroups (n = 610).

Competing interests

MH and PW are employees of PatientsLikeMe, and own stock or stock options in the company. The PatientsLikeMe research and development team has received research support from pharmaceutical companies and private foundations. VTT, VMM, CB and PR have no financial interests in PatientsLikeMe and no other conflicts of interest to disclose.

Authors' contribution

VTT, MH, VMM, CB, PW, and PR conceived and designed the study. MH and PW acquired the data. VTT and CB analyzed and interpreted the data. VTT drafted the manuscript. MH, VMM, CB, PW, and PR critically revised the manuscript for important intellectual content. PR and PW provided administrative, technical, and material support. All authors saw and approved the final manuscript. PR is the guarantor, had full access to the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgments

We thank Laura Smales (BioMedEditing), ADT International (Paris) and ISInfoglobal support (Barcelona) for editing and translation work, and Shimon Rura for help with data collection.

Funding

This study was funded by the French Health Ministry (PHRC AOM13127).

Author details

¹Department of General Medicine, Paris Diderot University, Paris, France. ²METHODS Team, Epidemiology and Statistics Sorbonne Paris Cité Research Centre, UMR 1153, INSERM, 1 place du Parvis Notre-Dame, Paris 75181, France. ³Centre d'Épidémiologie Clinique, Hôpital Hôtel-Dieu, Assistance Publique-Hôpitaux de Paris, Paris, France. ⁴PatientsLikeMe, Inc., Cambridge, Massachusetts, USA. ⁵Division of Health Care and Policy Research, Department of Health Sciences Research and Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, USA. ⁶Paris Descartes University, Paris, France. ⁷Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA.

Received: 7 February 2014 Accepted: 12 June 2014

Published: 2 July 2014

References

1. May C, Montori VM, Mair FS: We need minimally disruptive medicine. *BMJ* 2009, **339**:b2803.
2. Eton DT, Ramalho de Oliveira D, Egginton JS, Ridgeway JL, Odell L, May CR, Montori VM: Building a measurement framework of burden of treatment in complex patients with chronic conditions: a qualitative study. *Patient Relat Outcome Meas* 2012, **3**:39–49.
3. Eton DT, Elraiyah TA, Yost KJ, Ridgeway JL, Johnson A, Egginton JS, Mullan RJ, Murad MH, Erwin PJ, Montori VM: A systematic review of patient-reported measures of burden of treatment in three chronic diseases. *Patient Relat Outcome Meas* 2013, **4**:7–20.
4. Jowsey T, Yen L, PM W: Time spent on health related activities associated with chronic illness: a scoping literature review. *BMC Public Health* 2012, **12**:1044.
5. Boyd CM, Darer J, Boulton C, Fried LP, Boulton L, Wu AW: Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 2005, **294**:716–724.
6. Guthrie B, Payne K, Alderson P, McMurdo ME, Mercer SW: Adapting clinical guidelines to take account of multimorbidity. *BMJ* 2012, **345**:e6341.
7. Montori VM, Brito JP, Murad MH: The optimal practice of evidence-based medicine: incorporating patient preferences in practice guidelines. *JAMA* 2013, **310**:2503–2504.
8. Bohlen K, Scoville E, Shippee ND, May CR, Montori VM: Overwhelmed patients: a videographic analysis of how patients with type 2 diabetes and clinicians articulate and address treatment burden during clinical encounters. *Diabetes Care* 2012, **35**:47–49.
9. Tran VT, Montori VM, Eton DT, Baruch D, Falissard B, Ravaut P: Development and description of measurement properties of an instrument to assess Treatment Burden among patients with multiple chronic conditions. *BMC Med* 2012, **10**:68.
10. Eton DT, Oliveira DR, Egginton J, Mair FS, May C, Montori VM: Understanding the burden of treatment in patients with multiple chronic conditions: Evidence from exploratory interviews. *Qual Life Res* 2010, **19**:1673.
11. McKenna SP: Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. *BMC Med* 2011, **9**:86.
12. Acquadro C, Conway K, Hareendran A, Aaronson N: Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health* 2008, **11**:509–521.
13. Beaton DE, Bombardier C, Guillemin F, Ferraz MB: Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000, **25**:3186–3191.

14. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P: Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005, **8**:94–104.
15. Gallacher K, May CR, Montori VM, Mair FS: Understanding patients' experiences of treatment burden in chronic heart failure using normalization process theory. *Ann Fam Med* 2011, **9**:235–243.
16. Gallacher K, Morrison D, Jani B, Macdonald S, May CR, Montori VM, Erwin PJ, Batty GD, Eton DT, Langhorne P, Mair FS: Uncovering treatment burden as a key concept for stroke care: a systematic review of qualitative research. *PLoS Med* 2013, **10**:e1001473.
17. Sav A, Kendall E, McMillan SS, Kelly F, Whitty JA, King MA, Wheeler AJ: You say treatment, I say hard work: treatment burden among people with chronic illness and their carers in Australia. *Health Soc Care Community* 2013, **21**:665–674.
18. The Open Research Exchange. [http://www.openresearchexchange.com]
19. Wicks P, Heywood B, Heywood J: Online platform to accelerate patient involvement in open instrument development. *Qual Life Res* 2013, **22**:55.
20. PatientsLikeMe. [http://www.patientslikeme.com]
21. Cronbach L: Coefficient alpha and the internal structure of tests. *Psychometrika* 1951, **16**:297–334.
22. Nunnally J, Bernstein I: *Psychometric theory*. 3rd edition. New York: McGraw-Hill; 1994.
23. Slawsky K, Massagli M, Wicks P: PRM30 a comparison of the PatientsLikeMe Quality of Life Questionnaire (PLMQOL) with the RAND SF-36. *Value Health* 2011, **14**:A426.
24. Morisky DE, Ang A, Krousel-Wood M, Ward HJ: Predictive validity of a medication adherence measure in an outpatient setting. *J Clin Hypertens (Greenwich)* 2008, **10**:348–354.
25. Morisky DE, Green LW, Levine DM: Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care* 1986, **24**:67–74.
26. Diederichs C, Berger K, Bartels DB: The measurement of multiple chronic diseases—a systematic review on existing multimorbidity indices. *J Gerontol A BiolSci Med Sci* 2011, **66**:301–311.
27. Juniper E, Guyatt G, Jaeschke R: How to develop and validate a new health-related quality of life instrument. In *Quality of life and pharmacoeconomics in clinical trials*. Edited by Spilker B. Philadelphia - New York: Lippincott - Raven; 1996:49–56.
28. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979, **86**:420–428.
29. Streiner D, Norman G: *Health Measurement Scales: a practical guide to their development and use, Fourth Edition edn*. Oxford: Oxford University Press; 2008.
30. Bland JM, Altman DG: Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999, **8**:135–160.
31. The R Project for Statistical Computing. [http://www.R-project.org]
32. Ubel PA, Abernethy AP, Zafar SY: Full disclosure—out-of-pocket costs as side effects. *N Engl J Med* 2013, **369**:1484–1486.
33. Ekman A, Litton JE: New times, new needs; e-epidemiology. *Eur J Epidemiol* 2007, **22**:285–292.
34. Bove R, Secor E, Healy BC, Musallam A, Vaughan T, Glanz BI, Greeke E, Weiner HL, Chitnis T, Wicks P, De Jager PL: Evaluation of an online platform for multiple sclerosis research: patient description, validation of severity scale, and exploration of BMI effects on disease course. *PLoS One* 2013, **8**:e59707.
35. Wicks P, Sulham KA, Gnanasakthy A: Quality of life in organ transplant recipients participating in an online transplant community. *Patient* 2014, **7**:73–84.
36. De la Loge C, Keininger D, Isojärvi J, Massagli MP, Wicks P: Characteristics of users of the epilepsy community of PatientsLikeMe.com and comparison with a representative claims database. In *63rd Annual Meeting of the American Academy of Neurology*. Honolulu, USA; 2011.
37. Bowden A, Fox-Rushby JA: A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Soc Sci Med* 2003, **57**:1289–1306.

doi:10.1186/1741-7015-12-109

Cite this article as: Tran et al.: Adaptation and validation of the Treatment Burden Questionnaire (TBO) in English using an internet platform. *BMC Medicine* 2014 **12**:109.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Appendix 1: Demographic and clinical characteristics of patients included in the pretest of the English Treatment Burden Questionnaire (TBQ) (n=200).

Characteristic	Value
Age, yr	49.6 (12.5)
Female sex – no. (%)	141 (70.5%)
Country of residence – no. (%)	
United States of America	142 (71.0%)
Canada	19 (9.5%)
United Kingdom	20 (10.0%)
Australia and New Zealand	8 (4.0%)
Other	12 (6.0%)
No. of tablets and pills/day	9.7 (7.9)
No. of injections/week	0.7 (3.6)
No. of drug administrations/day	3.1 (1.8)
No. of different doctors the patient regularly sees)	3.3 (2.7)
No. of appointments/month	3.1 (4.2)
No. of hospitalizations/year	0.7 (2.0)
Presence of an informal caregiver – no. (%)	116 (58.0%)
Most frequent location for medical consultations – no. (%)	
Public hospital	17 (8.5%)
Private hospital	10 (5.0%)
General practitioner's clinic	94 (47.0%)
Specialist's clinic	79 (39.5%)
Duration of oldest chronic condition – no. (%)	
< 5 years	33 (16.5%)
5–10 years	31 (15.5%)
> 10 years	136 (68.0%)
No. of different chronic conditions	2.9 (2.1)
Conditions – no. (%)*	
Diabetes	20 (10.0%)
Other endocrine disorders	36 (18.0%)
Lung diseases	29 (14.5%)
High blood pressure	48 (24.0%)
Heart diseases	15 (7.5%)
Kidney diseases	12 (6.0%)
Gastro intestinal diseases	41 (20.5%)
Stroke or cerebrovascular diseases	9 (4.5%)
Neurologic diseases	110 (55.0%)
Rheumatologic diseases	64 (32.0%)
Cancer or malignant blood diseases	6 (3.0%)
Psychiatric diseases	95 (47.5%)
Vision problems	35 (17.5%)
Hearing problems	18 (9.0%)
Skin diseases	20 (10.0%)
Infectious diseases	3 (1.5%)
Fibromyalgia	22 (11.0%)

Data are mean (SD) unless indicated. *A given patient can have multiple conditions

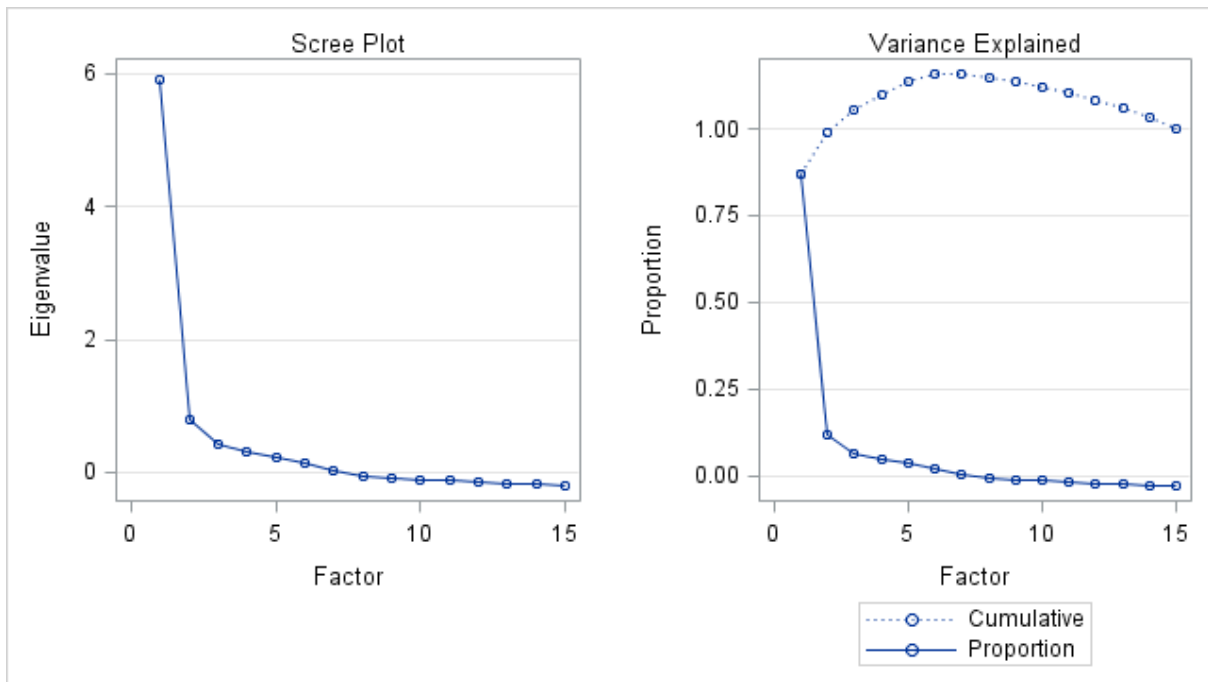
Appendix 2: Item properties of the TBQ (n=610 patients).

Item	Response rate No. (%)	Item “Does not apply” No. (%)	“No burden” (score=0) No. (%)	“High burden” (score >7) No. (%)
The taste, shape or size of your tablets and/or the annoyances caused by your injections (e.g., pain, bleeding, bruising or scars)	607 (99.5%)	44 (7.2%)	217 (35.6%)	95 (15.6%)
The number of times you should take your medication daily	606 (99.3%)	23 (3.8%)	184 (30.2%)	55 (9.0%)
The efforts you make not to forget to take your medications (e.g., managing your treatment when you are away from home, preparing and using pillboxes)	606 (99.3%)	19 (3.1%)	182 (29.8%)	85 (13.9%)
The necessary precautions when taking your medication (e.g., taking them at specific times of the day or meals, not being able to do certain things after taking medications such as driving or lying down)	607 (99.5%)	28 (4.6%)	171 (28.0%)	115 (18.8%)
Lab tests and other exams (e.g., blood tests or radiology): frequency, time spent and associated nuisances or inconveniences	608 (99.7%)	24 (3.9%)	173 (28.4%)	98 (16.1%)
Self-monitoring (e.g., taking your blood pressure or checking your blood sugar): frequency, time spent and associated nuisances or inconveniences	604 (99.0%)	185 (30.3%)	214 (35.1%)	43 (7.0%)
Doctor visits and other appointments: frequency and time spent for these visits and difficulties finding healthcare providers	607 (99.5%)	12 (2.0%)	151 (24.7%)	152 (24.9%)
The difficulties you could have in your relationships with healthcare providers (e.g., feeling not listened to enough or not taken seriously)	606 (99.3%)	4 (0.6%)	171 (28.0%)	189 (31.0%)
Arranging medical appointments and/or transportation (doctors visits, lab tests and other exams) and reorganizing your schedule around these appointments	606 (99.3%)	9 (1.5%)	166 (27.2%)	151 (24.7%)
The administrative burden related to healthcare (e.g., all you have to do for hospitalizations, insurance reimbursements and/or obtaining social services)	607 (99.5%)	44 (7.2%)	147 (24.1%)	136 (22.3%)
The financial burden associated with your healthcare (e.g., out-of-pocket expenses or expenses not covered by insurance)	608 (99.7%)	31 (5.1%)	101 (16.6%)	263 (43.1%)
The burden related to dietary changes (e.g., avoiding certain foods or alcohol, having to quit smoking)	608 (99.7%)	45 (7.4%)	172 (28.2%)	152 (24.9%)
The burden related to doctors' recommendations to practice physical activity (e.g., walking, jogging, swimming)	608 (99.7%)	35 (5.7%)	155 (25.4%)	153 (25.1%)
How does your healthcare impact your relationships with others (e.g., being dependent on others and feeling like a burden to them, being embarrassed to take your medications in public)	607 (99.5%)	18 (2.9%)	112 (18.4%)	240 (39.3%)
'The need for medical healthcare on a regular basis reminds me of my health problems'	607 (99.5%)	-	71 (11.6%)	295 (48.4%)

Appendix 3: Correlation of each item of the TBQ with the global score, omitting that item from the total (n=610).

TBQ Item	Spearman correlation coefficient
The taste, shape or size of your tablets and/or the annoyances caused by your injections (e.g., pain, bleeding, bruising or scars)	0.52 (p<0.0001)
The number of times you should take your medication daily	0.54 (p<0.0001)
The efforts you make not to forget to take your medications (e.g., managing your treatment when you are away from home, preparing and using pillboxes)	0.64 (p<0.0001)
The necessary precautions when taking your medication (e.g., taking them at specific times of the day or meals, not being able to do certain things after taking medications such as driving or lying down)	0.63 (p<0.0001)
Lab tests and other exams (e.g., blood tests or radiology): frequency, time spent and associated nuisances or inconveniences	0.64 (p<0.0001)
Self-monitoring (e.g., taking your blood pressure or checking your blood sugar): frequency, time spent and associated nuisances or inconveniences	0.47 (p<0.0001)
Doctor visits and other appointments: frequency and time spent for these visits and difficulties finding healthcare providers	0.70 (p<0.0001)
The difficulties you could have in your relationships with healthcare providers (e.g., feeling not listened to enough or not taken seriously)	0.58 (p<0.0001)
Arranging medical appointments and/or transportation (doctors visits, lab tests and other exams) and reorganizing your schedule around these appointments	0.71 (p<0.0001)
The administrative burden related to healthcare (e.g., all you have to do for hospitalizations, insurance reimbursements and/or obtaining social services)	0.58 (p<0.0001)
The financial burden associated with your healthcare (e.g., out-of-pocket expenses or expenses not covered by insurance)	0.54 (p<0.0001)
The burden related to dietary changes (e.g., avoiding certain foods or alcohol, having to quit smoking)	0.57 (p<0.0001)
The burden related to doctors' recommendations to practice physical activity (e.g., walking, jogging, swimming)	0.56 (p<0.0001)
How does your healthcare impact your relationships with others (e.g., being dependent on others and feeling like a burden to them, being embarrassed to take your medications in public)	0.63 (p<0.0001)
'The need for medical healthcare on a regular basis reminds me of my health problems'	0.50 (p<0.0001)

Appendix 4: Scree plots and Eigenvalues for the correlation matrix of the factor analysis (n=610). “Does not apply” was considered the lowest possible score (0).



Factor	Eigenvalue	Proportion of variance explained	Cumulative proportion of variance explained
1	5.91270808	0.8716	0.8716
2	0.80393808	0.1185	0.9901

Appendix 5: Correlation between variables and TBQ global score (n=610).

Clinical variables	Correlation with TBQ global score*	TBQ global score Mean (SD)	P value
Educational level High school diploma or below (n=123) College (n=478)	-	59.1 (32.4) 51.8 (31.2)	0.02
No. of different chronic conditions 1 (n=181) 2-3 (n=234) >4 (n=195)	-	44.3 (29.1) 49.7 (29.0) 65.4 (33.0)	<0.0001
No. of tablets and pills/day	0.20		<0.0001
No. of injections/week	0.11		0.02
No. of drug administration/day	0.25		<0.0001
No. of different doctors the patient regularly sees	0.21		<0.0001
No. of appointments/month	0.25		<0.0001
No. of hospitalization/year	0.11		0.01

*Correlations were assessed by Spearman's correlation coefficient. Comparisons between groups involved ANOVA.

Appendix 6: Reliability testing with test–retest method (n=282).

TBQ Item	ICC (95% CI)
The taste, shape or size of your tablets and/or the annoyances caused by your injections (e.g., pain, bleeding, bruising or scars)	0.57 (0.47-0.66)
The number of times you should take your medication daily	0.52 (0.40-0.62)
The efforts you make not to forget to take your medications (e.g., managing your treatment when you are away from home, preparing and using pillboxes)	0.58 (0.49-0.67)
The necessary precautions when taking your medication (e.g., taking them at specific times of the day or meals, not being able to do certain things after taking medications such as driving or lying down)	0.51 (0.40-0.60)
Lab tests and other exams (e.g., blood tests or radiology): frequency, time spent and associated nuisances or inconveniences	0.62 (0.51-0.70)
Self-monitoring (e.g., taking your blood pressure or checking your blood sugar): frequency, time spent and associated nuisances or inconveniences	0.57 (0.44-0.70)
Doctor visits and other appointments: frequency and time spent for these visits and difficulties finding healthcare providers	0.65 (0.57-0.73)
The difficulties you could have in your relationships with healthcare providers (e.g., feeling not listened to enough or not taken seriously)	0.70 (0.62-0.77)
Arranging medical appointments and/or transportation (doctors visits, lab tests and other exams) and reorganizing your schedule around these appointments	0.65 (0.56-0.72)
The administrative burden related to healthcare (e.g., all you have to do for hospitalizations, insurance reimbursements and/or obtaining social services)	0.59 (0.48-0.67)
The financial burden associated with your healthcare (e.g., out-of-pocket expenses or expenses not covered by insurance)	0.70 (0.61-0.76)
The burden related to dietary changes (e.g., avoiding certain foods or alcohol, having to quit smoking)	0.62 (0.52-0.70)
The burden related to doctors' recommendations to practice physical activity (e.g., walking, jogging, swimming)	0.58 (0.47-0.66)
How does your healthcare impact your relationships with others (e.g., being dependent on others and feeling like a burden to them, being embarrassed to take your medications in public)	0.70 (0.61-0.76)
The need for medical healthcare on a regular basis reminds me of my health problems	0.56 (0.45-0.65)
Global score*	0.77 (0.70-0.82)

Intraclass correlation coefficient 95% confidence interval was calculated by a bootstrap method.

Missing answers are considered the lowest possible score (0). *Global score is the sum of all items scores with “Does not apply” considered the lowest possible score (0).

Appendix 7: TBQ global score by different subgroups (n=610).

Subgroup	Global TBQ score Mean (SD)
Location of most frequent medical consultations	
Public hospital (n=61)	55.2 (33.1)
Private hospital (n=19)	52.9 (38.6)
General practitioner's clinic (n=286)	53.2 (30.3)
Specialist's clinic (n=220)	53.6 (32.0)
Types of treatment*	
Equipment (e.g., wheelchair, walking cane) (n=80)	69.0 (33.4)
Physical therapy (n=98)	62.7 (33.9)
Over the counter drugs (n=256)	59.1 (31.1)
Diet (n=101)	58.1 (32.6)
Exercises (n=159)	56.3 (33.3)
Complementary and alternative medicines (n=39)	55.8 (36.5)
Psychotherapy (n=219)	55.4 (33.3)
Supplements (e.g., vitamins) (n=399)	54.3 (31.5)
Prescription drugs (n=545)	54.0 (31.3)
Chronic conditions*	
Gastrointestinal diseases (n=128)	65.4 (32.5)
Skin diseases (n=68)	64.9 (30.8)
Fibromyalgia (n=77)	64.7 (32.2)
Lung diseases (n=90)	64.3 (35.0)
Rheumatologic diseases (n=201)	62.2 (31.7)
Psychiatric diseases (n=245)	61.3 (32.7)
Diabetes (n=42)	60.1 (35.6)
Other endocrine disorders (n=119)	57.8 (32.8)
Heart diseases (n=34)	57.8 (38.7)
Kidney diseases (n=37)	57.7 (36.8)
Vision problems (n=83)	57.5 (36.0)
Cancer or malignant blood diseases (n=30)	57.4 (36.3)
Hearing problems (n=48)	55.9 (30.4)
High blood pressure (n=153)	51.9 (31.3)
Neurologic diseases (n=270)	51.8 (30.1)
Infectious diseases (n=18)	51.2 (28.1)
Stroke or cerebrovascular diseases (n=17)	50.3 (39.2)

*A given patient can have multiple types of treatment and multiple conditions.

5. Décrire les composantes, facteurs aggravants et conséquences du fardeau du traitement

Nous venons de voir que les relations entre fardeau du traitement et système de soins rendaient complexe l'adaptation d'un unique outil dans des contextes, cultures et systèmes de soins différents. En effet, certains items ne seraient pertinents que dans certains contextes et pas d'autres. Dans ce cas, obtenir des mesures permettant de faire des comparaisons entre pays et contextes de soins, utilisables dans des essais cliniques internationaux peut être complexe. Pour une mesure du fardeau du traitement, cela est simplifié par l'hypothèse que le fardeau d'un aspect du traitement ne concernant pas le patient est nul. Autrement dit, si un patient ne pratique pas l'auto-surveillance de sa pression artérielle, alors nous pouvons considérer que l'auto-surveillance de la pression artérielle n'a pas d'impact sur sa qualité de vie.

A partir de cette hypothèse, nous pouvons supposer que pour construire un instrument global, permettant de comparer le fardeau du traitement entre plusieurs pays et contextes de soins différents –et de fait, utilisable dans des essais cliniques internationaux-, il faudrait que cet instrument couvre l'ensemble des aspects possibles du fardeau du traitement dans les contextes de soins cibles ; certains items non pertinents dans certains contextes seraient alors toujours « nuls ». En revanche, l'omission d'un item dans un contexte de soin pourrait conduire à une mesure « fausse » du concept. Par exemple, si nous n'avions pas ajouté l'item relatif au fardeau

financier du traitement dans la version anglaise du TBQ, la mesure du fardeau du traitement de ces patients aurait été incomplète.

Comme présenté dans le premier chapitre, pour définir les items pertinents d'un instrument et s'assurer de la validité de contenu, il est nécessaire de s'appuyer sur des études qualitatives menées auprès des patients cibles de l'instrument [89, 90]. Donc, pour créer un instrument « vraiment transculturel », utilisable dans plusieurs pays, il faut s'appuyer sur des recherches qualitatives ayant considéré les différences et similitudes entre les contextes de soins cibles et leur impact sur le fardeau. Depuis la publication de notre premier travail, plusieurs auteurs ont essayé de mieux définir les composantes du fardeau du traitement.

5.1. Etudes qualitatives sur le fardeau du traitement

Tout d'abord, Gallacher et al. a ré-analysé 47 interviews semi structurées de patients souffrant d'insuffisance cardiaque, au Royaume-Uni, portant sur le vécu de leur maladie et de leur traitement [71]. Pour cela, les auteurs ont utilisé le cadre théorique du « processus de normalisation » (en anglais : Normalization Process Theory). Ce cadre théorique vise à décrire comment les patients organisent, implémentent, maintiennent et transforment en routine les tâches qui leur sont imposées [91]. Ce travail a permis de définir une typologie du fardeau du traitement s'articulant autour de quatre concepts (**figure 5**) : 1) l'apprentissage des traitements et de leurs conséquences ; 2) la prise des traitements et l'implémentation des modifications des habitudes de vie ; 3) la surveillance des traitements ; et 4) les relations avec les autres. Ce travail était néanmoins limité par le fait qu'il s'agissait une ré-analyse de données déjà existantes, qui n'avaient pas été recueillies dans l'objectif de décrire le fardeau du

traitement. Ainsi, il était possible que certains aspects du concept n'aient pas été identifiés lors de cette étude.

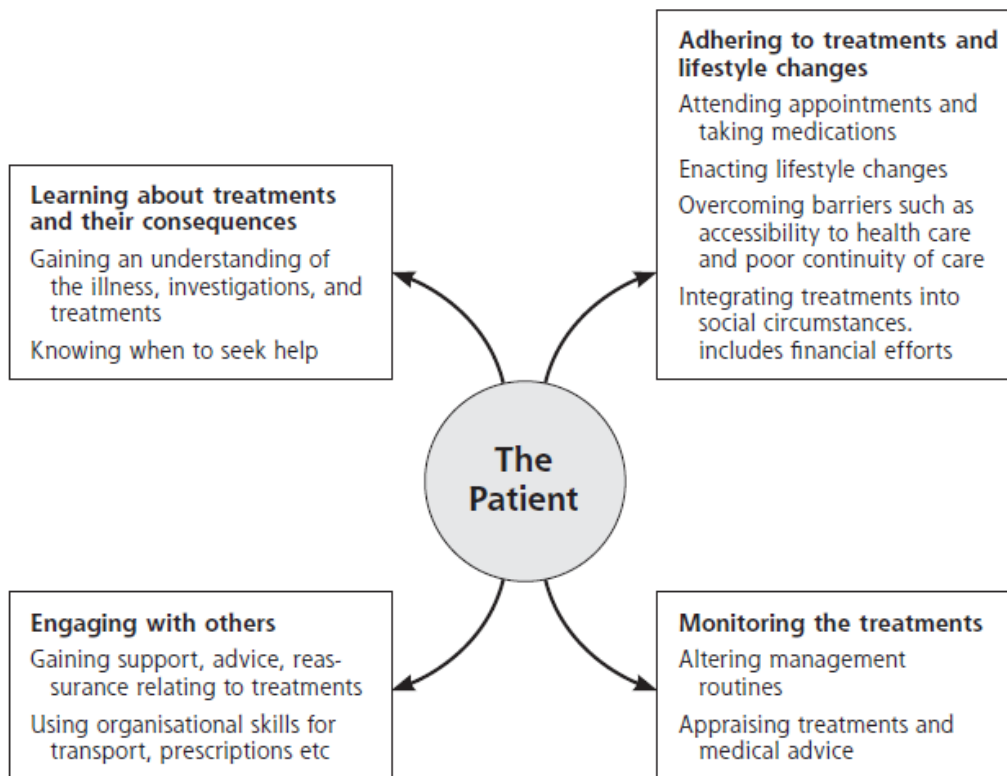


Figure 5: Typologie du fardeau du traitement chez des patients souffrant d'insuffisance cardiaque. Adapté de la source [71]

Dans un deuxième travail, Gallacher et al. a réalisé une méta-synthèse de 69 études qualitatives publiées entre 2000 et 2013 portant sur la perception par les patients de la prise en charge d'un accident vasculaire cérébral [70]. Les études incluses ont été ré-analysées à l'aide du cadre théorique du « processus de normalisation » décrit précédemment. Lors de ce travail, les auteurs ont retrouvé les quatre concepts déjà identifiés dans l'étude chez les patients souffrant d'insuffisance cardiaque. Bien qu'aucune étude ne porte spécifiquement sur le fardeau du traitement, l'inclusion de multiples études différentes a permis aux auteurs d'appréhender les différents aspects

de la prise en charge des patients pouvant représenter un problème pour eux. Notamment, les auteurs insistent sur le temps investi par les patients pour chercher, comprendre et appréhender les recommandations et prescriptions de leurs médecins vis-à-vis de la prise en charge de l'accident vasculaire cérébral. En effet, il apparaît dans cette étude que : 1) l'information prodiguée aux patients est pauvre, inadaptée ou peu compréhensible, 2) le temps que dédient les soignants à cette information est inadapté ; 3) cette information est souvent délivrée à un moment où les patients ne sont pas réceptifs à cette information. Outre le fait que les patients soient mal informés, les auteurs ont mis en évidence une relation entre le fardeau du traitement et l'organisation des soins mal adaptée au processus long de réadaptation dans les suites d'un accident vasculaire cérébral. Par exemple, les patients ont pu décrire leurs hospitalisations dédiées à la réadaptation, comme longues, ennuyeuses et non stimulantes. Les autres points mis en valeur par cette étude étaient les contraintes financières du traitement et les difficultés psychologiques que les patients pouvaient avoir pour rester motivés pendant les soins. Lors de cette étude, les auteurs reconnaissent d'une part que l'organisation des soins, variant entre les pays (voire entre les localités) joue un impact considérable sur le fardeau du traitement et que, d'autre part, il est nécessaire d'aborder le concept de fardeau du traitement de manière globale, en prenant en compte la multimorbidité.

Jani et al. a réalisé une méta-synthèse d'études qualitatives, similaire à celle de Gallacher et al., à partir de 16 études portant sur des patients avec une insuffisance cardiaque avancée [92]. Lors de cette analyse, les auteurs ont retrouvé des thèmes très similaires à ceux identifiés chez des patients ayant eu un accident vasculaire cérébral: 1) déficit d'information des patients, notamment en terme de pronostic ou de signes

cliniques « d'alerte » devant les amener à chercher de l'aide ; 2) le besoin de demander de l'aide à leurs proches ou à leur famille et les problèmes que cela pouvait poser; 3) les tâches imposées aux patients, (i.e. régime pauvre en sel, nombre important de traitements médicamenteux différents, etc.) et 4) les difficultés lors du suivi médical, décrit comme fragmenté et centré sur la maladie.

Bien que ces trois travaux permettent de mieux circonscrire le concept de fardeau du traitement, ils étaient limités car centrés sur une maladie spécifique. Dans son travail, Eton et al. a exploré le concept de fardeau du traitement sans restriction de maladie [47]. Pour cela, il s'est intéressé à 52 patients avec des traitements complexes affiliés à un programme de dispensation des traitements par des pharmaciens, à la Mayo Clinic, aux Etats-Unis. Ce travail a permis d'identifier trois grands thèmes autour du fardeau du traitement: 1) les tâches que les patients devaient faire ; 2) les stratégies, outils et compétences que les patients développaient pour accomplir ces tâches et gérer les problèmes liés à leur maladies ou traitements ; et 3) les facteurs aggravant le fardeau du traitement. Les facteurs aggravant identifiés dans cette étude étaient : les problèmes liés aux prises médicamenteuses ; les problèmes émotionnels impliquant la famille ou les proches ; l'impact du traitement sur le rôle social ou professionnel du patient ; les contraintes financières du traitement ; 5) les difficultés de compréhension des informations médicales et les obstacles systémiques à la prise en charge médicale. A notre connaissance, cette étude était la première dont l'objectif était d'identifier les composantes du fardeau du traitement. Elle était cependant limitée par le recrutement de ces patients, ne provenant que d'un seul centre, dans un pays donné.

De son côté, Sav et al. a réalisé une étude auprès de 97 participants (48 lors d'interviews face-à-face et 47 lors d'entretiens téléphoniques) recrutés dans quatre régions d'Australie. Les participants étaient soit des patients souffrant de maladies chroniques (définies comme une affection de santé durant plus de 6 mois), soit les aidants informels de patients. Cette étude a mis en évidence quatre thèmes du fardeau du traitement : 1) les contraintes financières ; 2) le temps nécessaire aux patients pour faire tout ce qui leur était recommandé, incluant le temps nécessaire pour aller aux rendez-vous ; 3) les contraintes directement liées aux traitements ; et 4) les contraintes liées à des obstacles systémiques (i.e. mauvaises relations médecin-malade, problèmes liés aux systèmes de remboursement, etc.).

Chaque étude présentée a permis d'éclairer une facette du concept du fardeau du traitement, que ce soit dans un contexte particulier (pathologie spécifique chez Gallacher et al. et Jani et al. ou cadre géographique spécifique chez Eton et al. et Sav et al.). Même s'il existait un recouvrement des concepts mis en évidence dans ces études par les items du TBQ, leur caractère restreint empêchait une généralisation.

5.2. Décrire le fardeau du traitement dans plusieurs pays et contextes de soins différents

Nous avons donc décidé de réaliser une étude qualitative multinationale, dans des pays de niveaux socio-économiques proches. Cette étude avait un double objectif : 1) compléter les travaux précédemment cités en identifiant les différentes composantes du fardeau du traitement chez des patients avec des pathologies chroniques, recevant des traitements variés, dans des pays et des contextes de soins différents ; et 2) recueillir les mots et pensées des patients sur le fardeau de leur traitement

directement, dans plusieurs langues, afin d'aider au développement d'items d'un outil transculturel pour mesurer le fardeau du traitement.

5.2.1. Méthodes

Pour explorer les difficultés des patients à gérer leurs maladies chroniques dans leur vie de tous les jours, nous avons réalisé une étude qualitative à l'aide de questions ouvertes dans un outil en ligne. L'outil en ligne correspondait à un site internet décrivant le concept du fardeau du traitement et à un questionnaire en ligne disponible en 3 langues (Français, Anglais, Espagnol).

Nous avons recruté des patients souffrant d'au moins une maladie chronique (définie comme une maladie requérant des soins >6 mois) par trois méthodes: 1) invitation par des médecins partenaires ; 2) invitation par des associations de patients partenaires ; et 3) publicité sur des forums en ligne et des réseaux sociaux. Les patients ayant participé étaient encouragés à inviter leurs proches ou amis répondants aux critères d'éligibilité, selon une méthode « boule de neige ».

Les patients inclus répondaient à un questionnaire en ligne comprenant : 1) une première question ouverte, très globale, sur le fardeau du traitement, destinée à explorer les contraintes qu'il pouvait avoir pour se soigner, sans l'influencer ; 2) 16 questions portant sur les principaux thèmes du fardeau du traitement précédemment identifiés dans la littérature et dans nos travaux précédents: les prises médicamenteuses, les examens complémentaires, les visites médicales, les changements d'habitudes de vie, les problèmes d'organisation ou de transports, les tâches administratives, les contraintes financières, l'impact social du traitement et les problèmes liés à l'organisation du système de santé ; et 3) une dernière question, de

nouveau très ouverte, destinée à explorer d'autres aspects du fardeau du traitement non mentionnés dans les questions précédentes.

Les données qualitatives ont été analysées selon une méthode d'analyse de contenu dérivée de la théorie ancrée. Tout d'abord, deux investigateurs (VTT et CB) ont lu, de manière indépendante, les 200 premières réponses en Anglais et en Français afin d'identifier les idées, mots et pensées utilisés par les participants pour décrire leur fardeau du traitement. Lors de réunion de consensus, ils ont ensuite défini une première liste de thèmes pertinents pour les patients. Dans une deuxième étape, la liste de thèmes obtenue a été appliquée au reste des réponses des participants. Chaque réponse a été lue et codée par deux investigateurs –dont un parlant couramment la langue du participant-, de manière indépendante. Lors de réunions, les investigateurs comparaient et arrivaient à un consensus sur leurs codages. Lorsqu'une nouvelle idée émergeait, les investigateurs en discutaient et enrichissaient, si besoin la liste de thèmes. Dans une troisième étape, un investigateur (VTT) a relu toutes les contributions des patients pour évaluer la cohérence du codage.

Dans un deuxième temps, le corpus a été analysé à l'aide d'une analyse textuelle automatique selon la méthode de Reinert. Cette méthode étudie l'organisation interne d'un texte plutôt que le sens de celui-ci [93, 94]. Lors de cette méthode, le texte est fragmenté en segments composés d'une dizaine de mots délimités par de la ponctuation. La méthode groupe les segments semblables entre eux en fonction des mots qu'ils contiennent lors d'une classification hiérarchique descendante. L'analyse des mots contenus dans chacune des classes obtenues permet l'identification des principaux thèmes d'un texte.

5.2.2. Résultats

Entre juin 2013 et mars 2014 ; 5492 personnes se sont connectés sur le site internet (ce nombre inclut les participants, des patients non éligibles, des médecins, des chercheurs, des membres d'associations de patients et des visiteurs curieux) et 1345 (24%) se sont identifiés comme éligibles. Parmi ces patients, 1053 (78%) ont répondu aux questions ouvertes sur le fardeau du traitement. L'âge moyen des participants était de 46 (SD=14) ans ; 766 (73%) étaient des femmes. 671 (64%) participants résidaient en France ; 140 (13%) aux Etats-Unis ; 66 (6.3%) au Canada ; 56 (5.3%) au Royaume Uni ; et 30 (2.8 %) en Australie.

L'analyse qualitative a permis de classer les réponses des participants en 3 grandes catégories : 1) les tâches imposées aux patients par leurs maladies ou par le système de soin ; 2) les facteurs aggravant le fardeau du traitement associé à ces tâches ; et 3) les conséquences de ces tâches sur la vie des patients. La taxonomie complète est présente dans l'article présenté après.

5.2.2.1. Tâches imposées aux patients par leurs maladies ou par le système de soin

Ces tâches pouvaient avoir trait : 1) à la gestion des traitements (i.e. gérer ses stocks de médicaments, les renouveler, les avoir toujours sur soi, etc.); 2) aux modifications des habitudes de vie (i.e. devoir arrêter de consommer certains aliments, de l'alcool, devoir faire ou arrêter une activité physique donnée, etc.) ; 3) au suivi de la maladie ou des traitements (i.e. organiser les visites, gérer les transports pour se rendre aux rendez-vous, etc.) ; 4) aux tâches administratives ou 5) au fait de devoir apprendre sa maladie et/ou son traitement. Voici un exemple de tâche administrative imposée aux patients avec des maladies chroniques : *« Il y a toujours un tas de papiers à faire, ou des*

justificatifs à aller chercher à droite et à gauche. Les gens ne se rendent pas compte que quand on est malade ça nous coute beaucoup. Personne ne nous conseille sur nos droits ou pour nous aider à remplir les papiers.» (Femme de 39 ans, résidant en France ayant un Lymphome digestif B à grandes cellules).

5.2.2.2. Facteurs aggravant le fardeau du traitement

Le fardeau du traitement pouvait être exacerbé par : 1) la nature, temps requis et fréquence des tâches imposées aux patients ; 2) des facteurs structurels (i.e. distance entre lieux de vie et structures de soins, délais d'attente, facilité à obtenir des consultations en urgence, etc.) ; 3) des facteurs personnels (i.e. croyances dans les effets du traitements, relation médecin-malade, etc.) ; 4) des facteurs situationnels (i.e. lors de voyages, lorsque le patient n'est pas à domicile, etc.) ; et 5) des facteurs financiers. Voici un exemple de relation médecin-patient aggravant le fardeau du traitement des patients : *«on est mal informé par certains médecins, une fois je suis allé chez mon néphrologue (...), je lui ai demandé quand est-ce que j'allais guérir (...) et là il m'a dit que je ne guérirais jamais car il n'y a pas de traitement (...) et que je devrais faire de la dialyse et me faire greffer. (...). Ce mot greffe m'a fait fondre en larmes. J'imaginai le pire. J'ai dû aller sur internet pour trouver des réponses aux questions que je me posais, certains médecins devraient mieux expliquer à leur patient»* (Femme de 27 ans, résidant en Belgique, avec une hypertension artérielle et une maladie rénale).

5.2.2.3. Conséquences du fardeau du traitement

Le fardeau du traitement avait un impact : 1) sur l'observance des tâches imposées aux patient ; 2) sur la vie familiale, sociale et/ou professionnelle du patient (i.e. devoir s'expliquer sans cesse, devoir se cacher pour prendre ses traitements, devoir sacrifier son travail ou ses loisirs pour dédier du temps à sa santé, etc.) ; 3) sur la vie personnelle

et émotionnelle du patient (i.e. culpabilité de ne pas faire toutes les tâches imposées, etc.) ; et 4) financier. Voici un exemple de conséquence du fardeau du traitement sur la vie familiale et personnelle des patients : « *Cela bloque toute possibilité d'aménagements de temps de repos significatifs (partir en villégiature) dans la mesure où je donne la priorité à mes examens médicaux (sans compter les annulations ou reports de la part des praticiens, ce qui peut décaler toute une série d'examens). Ex: je n'ai pas pu prendre deux semaines de vacances consécutives hors de mon lieu d'habitation depuis 9 mois.* » (Homme de 44 ans, vivant en France, ayant eu un accident vasculaire cérébral).

5.2.2.4. Analyse textuelle automatique

Dans un second temps, nous avons analysé les réponses des participants à l'aide d'une analyse textuelle automatique. En Anglais, l'analyse a abouti à la formation de 11 classes contenant chacune au moins 100 segments. Chaque classe était composée. Les classes mettaient en opposition les tâches que les patients réalisaient eux-mêmes, en général, directement liées aux soins (i.e. organiser les rendez-vous, les renouvellements, gérer les traitements, les effets secondaires) et les tâches impliquant ou ayant un impact sur les autres, c'est-à-dire ayant un impact sur le « monde extra médical » des patients (i.e. modification des habitudes de vie, relations avec les proches, aspects financiers). Les classes retrouvées étaient similaires dans les autres langues et lorsque l'on ne considérait que la première question ouverte.

5.2.3. Discussion

Notre étude a permis de développer la première taxonomie du fardeau du traitement chez des patients souffrant de maladies chroniques. Le fardeau du traitement dépend de l'intégration des tâches imposées par la maladie et/ou le traitement du patient. L'impact sur la qualité de vie du patient de chacune de ces tâches dépend des capacités

du patient, de l'environnement (i.e. système de soins, relations familiales et sociales, mais également circonstances particulières telles que les voyages, la grossesse, etc.). L'ensemble des tâches imposées aux patients ont un impact leur vie familiale, sociale, professionnelle. Lorsque les capacités du patient sont dépassées, alors le risque d'inobservance du traitement augmente.

C'est la première étude à avoir considéré une population de patients avec des maladies chroniques multinationale, sans restriction de lieu, maladie ou traitement. Nous avons ainsi pu appréhender les différences et similarités de la perception du fardeau du traitement, par des patients vivant dans des contextes différents, dans des pays différents avec des systèmes de soins différents. L'analyse des données a combiné deux méthodes, indépendantes –dont une automatique–, améliorant ainsi la crédibilité de ce travail. Nos résultats ont permis de compléter et préciser les modèles présentés précédemment par Eton et al., Gallacher et al. et Sav et al. En particulier, dans cette étude, nous avons pu identifier et décrire les facteurs aggravant le fardeau du traitement : mauvaise communication avec leurs médecins responsables de « fausses croyances » par rapport à leurs traitements, manque d'information de la part des soignants, mauvaise accessibilité des ressources de soins, etc. Ces différents points pourraient être la cible d'interventions destinées à réduire le fardeau du traitement et ainsi améliorer l'observance des patients et leur qualité de vie.

Par ailleurs, cette étude était la première à utiliser un outil internet pour recueillir des données qualitatives, simultanément, dans plusieurs pays différents. Nous avons donc pu démontrer la faisabilité d'une telle approche. En particulier, 78% des patients ayant « ouvert le questionnaire en ligne » sont allés jusqu'au bout et ont répondu aux

questions ouvertes sur le fardeau du traitement. Les avantages d'une telle méthode sont : 1) la disponibilité instantanée des réponses des patients, sans besoin de retranscription, comme dans la plupart des études qualitatives ; et 2) le faible coût de l'étude. En effet, recruter des patients géographiquement distants, dans plusieurs pays différents, aurait été complexe et coûteux ; nécessitant des investigateurs dans chaque pays d'étude.

Les limites de ce travail avaient trait aux limites de la méthode. En effet, l'utilisation d'un questionnaire en ligne n'autorisait pas les investigateurs à adapter son questionnement ou l'entretien aux réponses obtenues. De fait, les contributions de chaque patient sont significativement moins riches que si nous avions pu les interroger lors d'interviews « face à face ». Cependant, ceci était une contrepartie au fait que nous avons pu recruter un grand nombre de patients divers, permettant de consolider notre théorie et notre taxonomie du fardeau du traitement dans des groupes de patients différents tant sur le plan des maladies présentées, que des soins reçus et que des systèmes de soins dont ils faisaient partie. Une autre limite de cette étude venait du fait que nous n'avons recruté que des patients utilisant internet. De fait, nous avons probablement recruté des patients plus jeunes, d'un niveau d'éducation plus élevé et « doués » avec les ordinateurs [95]. Ces patients sont connus pour avoir de meilleures capacités pour répondre aux demandes imposées par l'environnement ; de fait l'impact du fardeau sur leur qualité de vie était probablement sous estimée, représentant une sorte de « cas idéal ». Finalement, les fréquences de chaque thème de cette étude ne correspondaient pas aux fréquences des différents aspects du fardeau du traitement présentés, dans cette population : les patients pouvaient ne pas mentionner certaines

contraintes qu'ils avaient à cause de problèmes de compréhension, de mémorisation, ou de biais de désirabilité sociale.

Cette étude nous a permis, à la fois, de développer une taxonomie du fardeau du traitement chez des patients souffrant de maladies chroniques utilisable quel que soit le contexte de maladie, traitement et système de soins du patient, et d'évaluer la faisabilité d'une étude qualitative internationale utilisant un recueil de données en ligne.

5.2.4. Article

Tran VT, Barnes C, Montori VM, Falissard B, Ravaud P. Taxonomy of the burden of treatment: a multi-country Web-based qualitative study of patients with chronic conditions. *In preparation*

Taxonomy of the burden of treatment: a multi-country Web-based qualitative study of patients with chronic conditions

Viet-Thi Tran (MD)^{1,2,3}, Caroline Barnes (MSc)⁴, Victor M Montori (MD)⁵, Bruno Falissard (MD, PhD)^{6,7}, Philippe Ravaud (MD, PhD)^{2,3,4,8}

¹ Department of General Medicine, Paris Diderot University, Paris, France

² Centre de recherche en Épidémiologie et Statistiques, INSERM U1153, Paris, France

³ Centre d'Épidémiologie Clinique, Hôpital Hôtel-Dieu, Assistance Publique-Hôpitaux de Paris, Paris, France

⁴ Paris Descartes University, Paris, France

⁵ Division of Health Care and Policy Research, Department of Health Sciences Research and Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, USA

⁶ Paris Sud University, Paris, France

⁷ INSERM U669, Paris, France

⁸ Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA

Abstract (298 words)

Background: Management strategies for patients with chronic conditions are becoming increasingly complex, which may result in a burden of treatment for patients. To develop a Minimally Disruptive Medicine designed to reduce the burden of treatment, clinicians need to understand which healthcare tasks and aggravating factors may be responsible for this burden.

Objective: Describe and classify the components of the burden of treatment for patients with chronic conditions from a patient point of view.

Methods: We performed a multi-country qualitative study using an online survey. We used a purposive sampling strategy to select English-, French- and Spanish-speaking participants with different chronic conditions. Participants were recruited by physicians, patients' associations, advertisement on social media, and "snowballing". The answers were analyzed by 1) manual content analysis with a grounded theory approach, coded by two researchers and 2) automatic textual analysis by Reinert's method.

Results: Between 2013 and 2014, 1,053 participants from 34 different countries completed the online survey using 408,625 words. Results from both analyses were synthesized in a taxonomy of the burden of treatment, which described 1) the tasks imposed on patients by their diseases and by their healthcare system (e.g., medication management, lifestyle changes, follow-up, etc.); 2) the structural (e.g., access to healthcare resources, coordination between care providers), personal, situational and financial factors that aggravated the burden of treatment; and 3) patient-reported consequences of the burden (e.g., poor adherence to treatments, financial burden, impact on professional, family and social life, etc.). Our findings may not be applicable to patients with chronic conditions who differ from those who responded to our survey.

Conclusion: Our taxonomy of the burden of treatment, provided by patients with chronic conditions from different countries and settings, supports the development of tools to ascertain the burden of treatment and highlights potential targets for interventions to minimize it.

1. Introduction

Multimorbidity is becoming increasingly common: 42% of patients have one chronic condition and 23% have multiple chronic conditions [1]. Multimorbid patients and their clinicians may struggle to balance the benefits and risks of multiple recommended treatments [2]. Indeed, physicians may be tempted to focus on individual diseases and follow clinical practice guidelines dedicated to one condition, but this approach may lead to overtreatment and unintended consequences [3, 4]. For example, a physician following clinical practice guidelines could prescribe up to 12 medications for a patient with osteoporosis, osteoarthritis, diabetes mellitus, hypertension, and chronic obstructive pulmonary disease [5]. As a result, in addition to the burden of illness, patients are affected by the burden of treatment, defined as the impact of the “work of being a patient” on functioning and well-being [6]. This work includes drug management, self-monitoring, visits to the doctor, laboratory tests, lifestyle changes, and other actions that take place in addition to the other work patients and their caregivers must do as part of life [7, 8]. Coping with all these healthcare tasks requires a significant amount of additional time, effort, and cognitive work from patients and caregivers and is associated with poor adherence to therapeutic care, independent of illness [9, 10].

The burden of treatment depends on patients’ investment of time and effort in following their physicians’ advice and on their context (e.g. social or family structure, care delivery system, etc.). This fact limits the transfer of findings from previous qualitative studies of specific conditions [11, 12] or in specific countries [7, 13]. Indeed, patients with similar conditions and treatment regimens could have different burdens of treatment depending on their education, culture, beliefs, family and social support, financial capacities, formal and informal support resources available, and healthcare context.

In the present study, we aimed to explore, describe and classify the components and consequences of the burden of treatment for patients with at least one chronic condition, across multiple conditions, treatments, countries and settings.

2. Methods

We performed a qualitative study using open-ended questions in an online survey, to explore patients' experiences and difficulties in managing their chronic conditions in everyday life. The Internet tool consisted of a website describing the concept of burden of treatment and an online questionnaire in three different languages (English, French, and Spanish).

2.1. Sample and recruitment

We recruited adult participants (>18 years old) with at least one chronic condition (defined as a condition requiring healthcare for at least 6 months) in three different ways: 1) invitation by participating physicians, 2) invitation by participating patient associations, and 3) advertisement on popular online health forums and social media (**Appendix 1**). Participants who had been invited were encouraged to invite their relatives and friends, who had chronic conditions, to participate, by a "snowball" sampling method[14], which involves identifying an initial number of participants who serve as "seeds" to help identify peers, who in turn are asked to invite others and so forth[15]. This sampling strategy was not designed to be representative of the population of patients with chronic conditions but rather sought to select a broad range of participants likely to experience different burdens of treatment.

All participants gave electronic informed consent before participating in the study. The study was reviewed and approved by the Institutional Review Board (IRB) of Cochin Hospital in France (no. 00001072) and the Mayo Clinic IRB (Rochester, MN, USA).

2.2. Data collection

In addition to demographic and clinical information, we collected qualitative data about patient experiences in managing chronic conditions in everyday life with open-ended questions.

First, we identified in the literature different aspects that contributed to the burden of treatment: taking medicines, self-surveillance (e.g. patients taking their blood pressure or measuring their blood sugar themselves, etc.), laboratory tests, doctor visits, learning about conditions and treatments, need for organization, transportation, administrative tasks, financial costs of treatment, difficulties in following advice on diet and physical exercise, social impact of the treatment, and problems associated with health organization (i.e., insurance coverage, access to care close to home, health policies, etc.)[7, 8, 11, 12, 16-19].

Second, we developed a preliminary questionnaire consisting of four parts: 1) demographic and clinical information about the participant's conditions and treatments; 2) a broad open-ended question at the beginning of the questionnaire to elicit the participant's view of the burden of treatment; 3) 16 open-ended questions about the aspects identified from the literature; and 4) a broad-ended question at the end of the questionnaire to identify other aspects of the burden of treatment that could have an impact on patients' quality of life but had not been assessed in the previous questions.

Third, this preliminary questionnaire was reviewed by seven physicians with experience in care of chronic conditions and pilot-tested, in pen-and-paper form, with 44 patients with chronic conditions recruited in university hospitals in Paris in November 2012 to assess the

clarity and wording of questions and types of answers. The final questionnaire was implemented online (<https://www.burdenoftreatment.com>). The ease of use and clarity of the Internet version were assessed by six patients and two physicians. All texts were translated into English and Spanish by professional translators and assessed by four native-speaking patients and investigators.

2.3.Data analysis

Quantitative data are described with means (SD) for continuous variables and proportions for categorical variables.

Qualitative data were analyzed by content analysis and automatic textual analysis. First, we analyzed all participants' responses using content analysis[20] with a grounded theory approach [21], to identify, for each patient, components and consequences of the burden of treatment. This analysis involved three steps. In a first step, two investigators (VTT and CB) independently identified for the first 200 responses in French and English, "in vivo codes": literal terms used by participants to explain and describe their burden of treatment. During meetings, the investigators reached consensus on the initial codes and grouped them into an initial set of themes that seemed meaningful to participants. Consensus was informed by the investigators' previous works on the burden of treatment[8, 9], the literature[7, 8, 11, 12, 16-19] and their clinical experience. In a second step, this initial set of themes was used for analysis of the remaining responses: each participant's response was read by two investigators (at least one researcher native in the given language), who independently assigned data segments to each theme. Analyses involved untranslated data: participants' verbatim answers were read in their original language, but coding was in English. During frequent meetings, the investigators compared their analyses and reached consensus on coding. Whenever a new idea emerged, researchers discussed the idea, thereby refining and enriching the list of themes. In a third step, an investigator (VTT) re-read participants' contributions to assess consistency with coding scheme.

Second, we examined the combined text of responses from all participants, for each language, using Reinert's automatic textual analysis method [22-25]. This method is a meaning-blind automatic textual analysis relying on the assumption that a text contains a reciprocal relationship between words and their proximate environments. For example, a section of text about health would include words related to health, and words related to health would be indicative of sections of text about health. Therefore, general ideas of a text could be revealed through the internal organization of the text. Reinert's method involves four steps: 1) creation of a basic vocabulary dictionary by identifying lexical forms (i.e., nouns, verbs, adjectives and adverbs) contained in the corpus; 2) fragmentation of the corpus into small parts, or elemental context units (ECUs), consisting of approximately 10 to 15 words delimited by punctuation; 3) creation of a table of lexical forms and ECUs; and 4) partition of the table into classes, using hierarchical descendant classification to group sections of texts according to their similarity (presence or absence) regarding the words in the texts using Chi-square tests. Analysis of words for each class allows for identification of topics covered in a text. The number of clusters was arbitrarily defined to contain at least 100 sections of text. Automatic textual analysis involved use of IRaMuTeQ 0.6 alpha 3 (<http://www.iramuteq.org/>).

2.4. Creation of a taxonomy of the burden of treatment

We developed a taxonomy using methods described by Bradley et al [26]. In a first step, two investigators (VTT, CB) used the different themes identified during the manual analysis to delineate an initial classification. In a second step, they used results from the automatic textual analysis to refine the classification. Finally, several meetings were held between researchers to discuss the creation of the taxonomy in light of previous studies about the burden of treatment.

2.5. Relationships between burden of treatment and respondent characteristics

To understand the context of each component of the burden of treatment, we examined the relationships between the patient's statement of a specific burden and clinical variables. Logistic regression was used to summarize these relationships. Odd ratios (ORs) were adjusted for key confounding factors (age, sex, presence of multimorbidity and educational level). We performed sensitivity analyses by using 2 definitions of multimorbidity (2 or more and 3 or more chronic conditions). Indeed, multimorbidity defined as patients with >2 chronic conditions may lack specificity because of the high proportion of patients involved [27]. Some authors argued that using >3 disease entities would likely identify patients with greater health needs and would therefore be more useful to clinicians [28]. Analyses involved use of SAS 9.3 (SAS Institute, Cary, NC, USA).

3. Results

From June 22, 2013 to March 30, 2014, 5,492 people connected to the Internet tool, and 1,345 people (24%) identified themselves as eligible for the study; 1,267 (94%) completed the demographic and clinical part of the survey, and 1,053 (78%) answered the open-ended questions about the burden of treatment (**Figure 1**). The mean (SD) age was 46 (14) years and 766 (73%) were women (**Table 1**). In total, 671 patients resided in France (64%), 140 in the United States (13%), 66 in Canada (6.3%), 56 in the United Kingdom (5.3%), 34 in Spain (3.2%), 30 in Australia (2.8%) and 56 (5.3%) in a different country. Self-reported chronic conditions included rheumatologic diseases (33%), cancer (8%), and well-controlled psychiatric illnesses (13%). The mean (SD) number of chronic conditions was 2.4 (1.6, range 1–10). A total of 662 patients (63%) had two or more chronic conditions.

Answers to open-ended questions formed a corpus of 408,625 words in English (148,707 words), French (243,558 words) and Spanish (16,360 words). Mean (SD) length of patients' answers were 388 (359) words globally (maximum 2699) and 108 (120) words for the first question (maximum 1267).

Manual content qualitative analysis and automatic textual analysis provided coherent results and described a list of difficulties patients could have when performing healthcare-related tasks. These two analyses were synthesized in light of the literature to construct a taxonomy of the burden of treatment. This taxonomy described 1) the tasks imposed on patients by their diseases and by their healthcare system, 2) the factors that exacerbate the burden associated with these tasks, and 3) how these tasks affected patients' lives (**Figure 2**).

3.1.Manual content qualitative analysis

For clarity, results of the manual qualitative analysis are presented according to the final classification.

3.1.1. Healthcare tasks imposed on patients

Most patients acknowledged that the care associated with having a chronic condition imposed a number of extra tasks (**Appendix3**): *“Being sick also adds a lot of extra tasks, paperwork and appointments. To keep myself healthy, I miss out on a lot of things that people my age take for granted - working fulltime, cooking, showering every day, going out to socialize”* (25-year-old woman with Ehlers-Danlos Syndrome from the United Kingdom). These tasks could be classified as 1) managing medications, 2) lifestyle changes, 3) condition and treatment follow-up, 4) paperwork tasks, and 5) learning about the condition and treatment.

3.1.1.1. Managing medications

Managing medications encompassed all tasks related to 1) medication intake and necessary precautions before or during intake, 2) adaptation and planning of daily life to take medications, and 3) ensuring medication stock (e.g., refills, carrying medication at all times). Implementation of “simple” prescriptions could sometimes result in major investments of time, energy or cognitive effort from patients. For example, *“Doctors often forget the time it takes to do a treatment isn't just the treatment itself. A nebulizer may finish in 6 minutes but that doesn't include mixing the drug and cleaning and sterilizing the equipment. That bumps the time to 20 min and doing that up to 5 times a day is suddenly a large chunk of your day”* (28-year-old woman with a pulmonary and a gastrointestinal condition from the United Kingdom). In total, 676 participants (64%) expressed at least one burden related to management of medication.

3.1.1.2.Lifestyle changes

Lifestyle changes referred to the efforts patients made to follow their physicians’ recommendations to 1) avoid eating certain foods (including paying attention to ingredients or having to plan and prepare their own meals), drinking alcohol or smoking; or 2) practice or give up physical exercise. For most participants, lifestyle changes were considered the most difficult tasks. For example, *“This makes me feel socially isolated. I cannot share much food with others, have to bring my own food, must deal with the hassle of having to explain to others why I cannot eat whatever I want, have the stigma of feeling weird and different from other people”* (57-year-old woman with pulmonary, dermatologic and rheumatologic conditions from the United States). In our study, 594 participants (56%) expressed at least a burden related to lifestyle changes.

3.1.1.3.Condition and treatment follow-up

Condition and treatment follow-up referred to all doctor visits, lab tests and complementary exams and self-monitoring patients had to organize, schedule and attend. The burden associated with these tasks was closely intertwined with structural factors (e.g., coordination between healthcare providers, access to health resources). For example, *“For many appointments, you must leave time for: getting to the appointment, finding parking, waiting for the appointment, seeing the doctor, getting back home. That can easily wipe out a morning*

or an afternoon” (75-year-old woman with thyroid disease, high blood pressure and a rheumatologic condition from Canada). Overall, 527 participants (50%) expressed at least a burden associated with tasks of condition/treatment follow-up.

3.1.1.4. Paperwork tasks

Paperwork tasks required time and cognitive energy for patients. They had to learn, understand and deal with complex rules and requirements necessary for their care and reimbursements. For example, *“It is difficult to fill these kinds of forms by yourself as there is a special language/code - there are certain things you have to say and certain things you have to avoid saying. It is sometimes very difficult to get help from a social worker or someone else who knows how these things work”*(28-year-old woman with cystic fibrosis from the United Kingdom). A total of 304 participants (29%) expressed at least a burden associated with paperwork tasks.

3.1.1.5. Learning about and developing an understanding of the imposed tasks

Learning and developing an understanding of the imposed tasks was cited as a mandatory task for patients with chronic conditions. It involved making sense and keeping up to date of medical jargon and often-conflicting information from different sources (e.g., relatives, Internet, nurses, physicians) before implementing this knowledge into their lives. For example, *“When I first started to take all the different medications it was completely overwhelming, learning how each tablet interacted with another one, how some had to be taken with food, some 1 hour before food, some 1 hour after food, some not within 2 hours of another one, some tablets can only be taken 12 hours apart, it can seem ridiculous until you get into a routine”* (40-year-old woman with endocrine and renal disorders from Australia). In all, 133 participants (13%) expressed at least a burden associated with making sense of everything that was asked of them.

3.1.2. Factors that exacerbate the burden of treatment

The burden of healthcare tasks imposed on patients could be exacerbated by 1) the nature, time required and frequency of the imposed tasks; 2) structural factors; 3) personal factors; 4) situational factors; and 5) financial factors.

3.1.2.1. Nature, time required and frequency of healthcare tasks

The nature (e.g., size, taste of tablets, discomfort associated with injections or blood tests, side effects of medications, life-long treatment duration), time required, and number and frequency of tasks imposed on patients were frequently cited as an aggravating factor of the burden of treatment (**Appendix 4**). For example, *“It is slow to take tablets - cannot take too many at a time because of gagging/nausea. Also some have a very strong bitter taste - on days when my hands are not working well it is difficult to get them in, sip water and swallow quickly”* (39-year-old woman with an auto-immune condition from China). In our study, 641 participants (61%), 422 participants (40%) and 638 participants (60%) expressed that their burden of treatment was exacerbated by its nature, the time required and the frequency of the imposed tasks, respectively.

3.1.2.2. Structural factors

Structural factors depend on the healthcare system in the country of residence. They relate to 1) poor access to healthcare resources (e.g., medication is not available at the pharmacy, distance to healthcare providers, difficulties in obtaining test results); 2) lack of coordination between care providers, who often care for patients individually without integrating their care in a more global process; 3) problems directly related to health centers (e.g., wait lines, lack of parking space); 4) lack of meaningful research on their conditions or treatments (e.g. including both lack of available research on specific rare chronic conditions and/or lack of knowledge of research in providers); and 5) inadequate public coverage of their conditions or treatments (**Appendix 5**). For example, *“Trips to hospital at least every three months (are) burdensome because of distance and also parking charges. Annual reviews are difficult because there are multiple investigations required and these take place over multiple different appointments and sometimes at different hospitals in the same group. Organizing repeat prescriptions is difficult: GPs are monitored for their prescribing and most of my medication comes from my GP. GP is not allowed any flexibility over prescribing (...): this means multiple trips to GP and pharmacy every month”* (55-year-old woman with diabetes and a cardiac condition from the United Kingdom). Overall, 366 participants (35%) told us that their burden was exacerbated by poor access to healthcare resources, 100 (9.5%) commented on the lack of coordination between their healthcare providers, and 316 (30%) expressed difficulties with health care centers.

3.1.2.3. Personal factors

Personal factors encompassed all factors related to 1) patients' beliefs about their conditions and treatment (e.g., feeling that their treatment lacked efficacy or that they were dependant on their medications); 2) the difficulties patients could have in their interactions with others (e.g., patients' feeling that they were a burden on their relatives, patients' desire to hide their condition/treatment from others): *“I'm a burden for my family. I need more attention and care than my children”* (45-year-old man with hemi-paretic stroke from France); and 3) difficulties patients could have in their interactions with healthcare providers: *“I think it's ridiculous that I have to do follow-ups every month when my doctor doesn't appear to know much about my disease and definitely doesn't listen to me when I'm telling him how I'm doing and doesn't answer questions”* (33-year-old woman with cancer, a rheumatologic condition and depression from the United States) (**Appendix 6**). A total of 252 participants (24%) expressed that some beliefs about their conditions or treatment aggravated the burden associated with healthcare tasks; 402 (38%) said that some interactions with others could exacerbate their burden of treatment; and 230 (22%) participants commented on how difficulties in their relationships with healthcare providers added to their burden of treatment.

3.1.2.4. Situational factors

Situational factors encompassed out-of-routine situations. Most participants, including those who had reconfigured their lives with their treatment and set up daily routines, cited the difficulties they had in performing healthcare-related tasks when they were away from home and/or in presence of other people (**Appendix 7**). For example, *“Good grief (...) I'm away on a vacation right now, and I had to go out the afternoon before we left to go buy more pill*

organizers, because I realized I didn't have enough for the amount of time this trip was going to take” (46-year-old woman with a neurological condition from the United States). In our study, 431 participants (41%) expressed difficulties adapting to out-of-routine situations.

3.1.2.5. Financial factors

Financial factors could impede patients' ability to do everything that was asked of them. These factors often were both aggravating factors and consequences of the burden of treatment. For example, *“I do not have insurance. I pay for every blood test, medication, doctor visit and the failed radioactive iodine uptake out of my own pocket. It is outrageously expensive and a tremendous burden”* (38-year-old man with a thyroid disease from the United States).

3.1.3. Consequences of healthcare tasks imposed on patients on their daily lives

Consequences of the aforementioned workload of healthcare involved 1) poor adherence to the tasks imposed on patients; 2) impact on family, social and professional life; 3) personal and emotional impact; and 4) financial impact of healthcare (**Appendix 8**).

3.1.3.1. Poor adherence to healthcare tasks

Poor adherence associated with the burden of treatment could be classified as intentional and non-intentional non-adherence. Intentional non-adherence involved patients giving up some tasks asked of them (from medication intake to administrative tasks) because of the additional financial costs or because it was “too much”. This non-adherence was different from intentional non-adherence, which follows from having a different understanding of disease and treatment that reduces concern for treating the disease or increases concern about the treatment's safety or lack of efficacy. An example of intentional non-adherence due to treatment burden follows: *“There is stuff that I am SUPPOSED to do, and stuff that I actually DO. If I did everything I am SUPPOSED to do, my life would revolve around doctors and tests and such and there wouldn't be very much left for living my life. So I've made a bunch of choices”* (46-year-old woman with a mitochondrial disease from the United States). Non-intentional non-adherence refers to the difficulties patients had in remembering to take medications or attend appointments, especially with a multitude of imposed tasks. As a result, patients developed strategies not to forget taking their medications: 1) use of pillboxes, 2) telephone reminders, 3) calendar, 4) daily actions or rituals, and 5) help from relatives. In our study, 64 participants (6.1%) reported a form of intentional non-adherence and 392 (37%) expressed having problems not to forget to follow their treatments.

3.1.3.2. Impact on professional, social, family life and leisure activities

Impact on professional, social, family life and leisure activities was defined as consequences of healthcare on patients' capacity to 1) work, from the time taken by healthcare activities and the difficulties they had to arrange with their co-workers to keep up their work, to the judgment of colleagues or employers; 2) participate and live a family, social or couple life; and 3) spend free time doing what they wanted. For example, *“Friendships are difficult to maintain because you end up cancelling plans and feel guilty every time you cancel on a friend or a family member. I would never wish a chronic illness on anyone”* (37-year-old woman with an endocrine and a rheumatologic condition from Canada). In total, 555

participants (53%) mentioned a negative consequence of healthcare activities on their professional, social and family life.

3.1.3.3. Emotional impact

The emotional impact of the burden of treatment related to 1) the frustration patients could experience by not being able to do everything they could have done because of their treatments: “*Low cholesterol, gluten sensitive, low calorie -- I love food but feel like I'm never really eating –(...) I'm a prisoner to food.*” (64-year-old woman with high blood pressure, vision problems, cardiac and rheumatologic conditions from the United States); 2) the guilt patients could have when not following prescriptions or recommendations; and 3) regular healthcare, which reminded them that they had chronic conditions. “*Medication intakes remind us that we have a lifelong incurable condition even if we look normal. It's depressing*” (19-year-old woman with systemic lupus from France). Overall, 325 participants (31%) mentioned that their treatment was responsible for a negative emotional impact on their lives.

3.1.3.4. Financial impact

Financial impact of the treatment represented both direct costs of healthcare demands and indirect costs associated with treatment (e.g., losing a job, costs of specific foods). This burden greatly varied between contexts and was closely related to both structural factors and patients' ability to deal with paperwork to get reimbursed. As stated earlier, financial factors were both consequences and aggravating factors of the burden of treatment, for a vicious circle. For example, “*All my money goes on my health aside from basic bills. I do not buy treats, clothes, haircuts, toiletries, things for the house (...) Have to spend a lot of time and energy on budgeting and I delay treatment sometimes as I have to save up*” (37-year-old woman with pernicious anemia and vision problems from the United Kingdom). In our study, 225 participants (21%) mentioned a financial burden of their treatment.

3.2. Automatic textual analysis

Automatic textual analysis was performed for each language separately. For English, we obtained 11 classes with at least 100 elemental context units (ECUs), after descendant hierarchical classification. Dendograms and words strongly associated with each class (chi-square >50) are in **Figure 2 Appendix 9**.

The organization of text reflected the different tasks participants performed to take care of their health. Each class consisted of words describing the different aspects of each problem (e.g., blood-test, lab, visit, scan, self-monitor) and words highlighting the difficulties encountered (e.g., wait, time, rush, etc.).

During classification, we found a contrast between tasks patients often performed themselves directly related to healthcare (organize tests, appointments, refills, manage medications and cope with side effects) and tasks that involved or had an impact on others, reflecting “out-of-medical-world interactions” (lifestyle changes, relationships with family, friends, society, financial aspects of healthcare). Concerning the tasks patients directly related to healthcare, we found a distinction between medication and treatment management (side effects, strategies not to forget to take medications and organization during travel) (23% of the corpus) and “out-of-home” care (tests, doctor visits, refills) (28% of the corpus). Concerning tasks

involving others out of the medical world, we found a distinction between paperwork and financial burden (14% of the corpus), which were closely related, and lifestyle changes and relationships with family, friends and other people (34% of the corpus).

We found a similar organization of the corpus for “spontaneous answers” (i.e., answers to the first open-ended question). For the two other languages, we found a greater emphasis on self-monitoring for tasks patients performed themselves (**Appendix 10**). This finding may be explained in part by the greater prevalence of conditions in which self-monitoring was integral to care, such as diabetes among Spanish-speaking respondents.

3.2. Relationships between components of the burden of treatment and respondents’ characteristics

In our sample, participants with > 2 chronic conditions more often elicited problems related to drug intake (OR=1.7 [1.3-2.4]), paperwork tasks (OR=1.4 [1.1-1.9]) and time required for tasks (OR=1.4 [1.0-1.8]) than those with only one chronic condition. They more often felt like they were a burden to others (OR=1.5 [1.0-2.2]), elicited more problems in their relationships with their healthcare providers (OR=1.6 [1.2-2.3]), and expressed more adherence issues (OR=2.2 [1.2-4]). When considering multimorbidity as patients with >3 chronic conditions, we found fewer differences between proportion of components of the burden of treatment identified by participants with or without multimorbidity, except for lack of adherence (**Appendix 11**). Older patients (>50 years old) elicited less often problems related to their professional life (OR=0.34 [0.25-0.46]) (**Appendix 12**). Female participants and those with higher educational level (College) were more likely to elicit different components of the burden of treatment (**Appendix 13 and 14**).

4. Discussion

With data from our qualitative study, we developed a comprehensive taxonomy of the burden of treatment for patients with chronic conditions in different contexts and countries. Independent analyses by informed investigators and a meaning-blind automatic procedure provided coherent results. Whereas manual qualitative analysis allowed us to explore the variety of themes mentioned by participants, automatic analysis allowed us to better understand how patients could dichotomize their lives between what their life in hospitals or clinics and their lives outside of the medical world, the latter sometimes shrinking in favor of the first. In our study, patients mentioned a number of healthcare tasks imposed on them. The burden associated with these tasks could be aggravated by multiple personal, structural or situational factors and affect quality of life and adherence to treatments, especially when consequences of treatment were not immediately visible or when treatment required extensive lifestyle changes. To our knowledge, this is the first study to provide a comprehensive view of the components and consequences of the burden of treatment across multiple countries, settings and treatment context. Our findings fit the Cumulative Complexity Model [29] in that the burden of treatment perceived by patients was a complex phenomenon resulting from the combination of 1) the workload of healthcare imposed on patients; 2) patients’ capacities to integrate this workload of healthcare in their daily lives in terms of energy, time, money or support; and 3) the context, especially the structure of healthcare (i.e., travel distance for

physician visits, waiting times, coordination between healthcare providers, reimbursements, etc.) and social support from their families, relatives, and friends.

In the present study, we assessed the relationships between patients' statements of specific components of the burden of treatment and their characteristics. Multimorbid patients were more likely to express concerns about drug intake, time required for healthcare tasks and talked significantly more often about adherence problems. These results agree with those from observational studies of the workload of care of multimorbid patients [8, 30] and highlight the magnitude of what patients perform unbeknownst to their physicians [30, 31]. If physicians spend about 2 hours each year with diabetic patients, these patients spend approximately 870 hours managing the disease on their own [32]. Because this involvement of time and effort is not usually discussed in depth during medical consultations [33], physicians should use adequate tools to diagnose and assess the burden of treatment [8].

The strength of this study relies in our focus on reducing the researcher's impact on the analysis by involving multiple researchers in the manual content analysis, each researcher supplementing and contesting others' statements [34], and by performing two separate analysis, one automatic. Our results contribute to the body of literature describing the burden of treatment for patients with at least 1 chronic condition. Compared to the work of Gallacher et al. [11], we gathered comparatively more data from a larger number of diverse participants. Thus, our findings may have broader applicability to the general population of patients with chronic conditions. Differences between our two classifications result from conceptual choices: Gallacher et al. used the processes patients use to implement tasks in their everyday life as key domains, whereas we used the tasks themselves. Similar to Eton et al. [7] and Sav et al. [13], we used as key domains 1) the imposed tasks and 2) aggravating factors of the burden of treatment. However, contrary to previous works, we integrated consequences of the burden of treatment in the taxonomy. Indeed, our analyses showed that patients indicated all aspects of the burden of treatment, including consequences, as a whole.

Our taxonomy of the burden of treatment also compliments the work with the Treatment Burden Questionnaire (TBQ) [8, 9]. Themes identified in the present study overlapped with the items of the TBQ but offered more precision and details, especially concerning the consequences and aggravating factors of the burden of treatment. This finding was not unexpected as the TBQ was developed to offer physicians tools to identify overburdened patients, in practice. As a result, the tool was voluntarily short and concise. Our taxonomy of the burden of treatment offers solid qualitative bases for development of specific and context-dependent measures that should compliment generic measures of the burden of treatment, like the TBQ.

An original aspect of this study was to use an online survey to gather data from geographically distant participants. Previous studies aiming to identify key concepts of a given topic in multiple countries were complex because they involved participation of researchers in each country [35]. In this study, we demonstrated the feasibility of using the Internet to gain insight into the experiences of patients directly and generate qualitative data simultaneously in multiple countries at low cost.

This study has some limitations. First, we used an online questionnaire with questions in a rigid order with no adaptation to prior responses or in-depth probing for more specific information in response to patients' statements. As a result, each patient's contribution may be less rich than what could have resulted from an interactive face-to-face interview[36]. Thus, this situation represents a trade-off from the gain in diversity of respondents. Second, the questionnaire could have influenced participant responding (e.g. by providing a starting point rather than allowing the patient to truly respond from the "ground" up). However, we found the exact same categories when considering only the first broad open-ended questions with subsequent questions allowing us to precise more fine grain themes. Third, the applicability of our findings may be limited given that our patients are not representative of the population of patients with at least one chronic condition in those countries (especially in Spanish-speaking countries, where we recruited only a small number of participants) or in the world. We only recruited patients using the Internet, which may have selected a population of respondents who were younger, more educated and "computer savvy"[37]. For example, in our study, only one patient mentioned language barriers as a burden of treatment, although it is an important barrier to healthcare access[38, 39]. Similarly, we did not account for income or insurance coverage, which are likely associated with patient-reported burden of treatment. Given the proportion of patients with high educational level, participants with lower socioeconomic were likely underrepresented in our sample. Because of the link between socioeconomic status and multimorbidity, the resource burdens and demands on patients' financial capacity may have been underrepresented, for a sort of "best-case" scenario. Finally, the variation in how patients answered the questionnaire prevents from drawing conclusions other than the diversity of the components, aggravating factors and consequences of the burden of treatment. Some patients may have not mentioned some of the difficulties with their treatments because of poor understanding of questions, recall problems and/or social desirability bias [40]. For example, participants with lower educational level were less likely to elicit components of the burden of treatment than those with higher educational level.

Our findings have several implications. First, our results may help clinicians better understand and identify patients who are overwhelmed by their treatments. In a previous study, we have shown that physicians fail to assess the burden of treatment of their patients[8], partly because it expresses a patient experience that is not shared in depth during consultations[33]. Second, our study points to the dire need to redesign guidelines to take into account multimorbidity[4, 41]. We found that multimorbid patients have more difficulties due to fragmentation of care; 60% of patients had 2 or more chronic conditions. They expressed a significantly greater number of difficulties performing healthcare tasks, independent of structural factors, age and country. Healthcare should be integrated and coherent: every therapeutic intervention imposed on patients should be carefully weighed in terms of clinical benefit, interaction with other treatments, possible harms, and potential burdens. Such consideration could result in the prioritization of tasks and a net reduction in healthcare tasks imposed on patients. Third, on a research level, new interventions should be designed to mitigate the aggravating factors identified, to improve patient adherence to treatment and to reduce the unintended negative impact of such treatments, through the burden they impose on patients' capacity, on their quality of life. Finally, on a structural level, this study highlights the need to change the

paradigm of care for patients with chronic conditions and end fractured care focused on individual conditions. Treatment objectives should not be based solely on attaining specific goals in specific diseases but also on mitigating the impact of conditions and treatments on patients' professional, family and social lives[42], for minimally disruptive medicine.

Conclusions

Data from our qualitative study of patients with different chronic conditions, in different contexts and countries provides a comprehensive taxonomy of the burden of treatment for such patients. Results could inform the development of cross-cultural instruments to assess the burden of treatment for patients with chronic conditions and new interventions to reduce the burden of treatment, ultimately moving towards minimally disruptive medicine [6].

References

1. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B: **Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study.** *Lancet* 2012, **380**(9836):37-43.
2. Bower P, Macdonald W, Harkness E, Gask L, Kendrick T, Valderas JM, Dickens C, Blakeman T, Sibbald B: **Multimorbidity, service organization and clinical decision making in primary care: a qualitative study.** *Fam Pract* 2011, **28**(5):579-587.
3. Tinetti ME, Fried T: **The end of the disease era.** *Am J Med* 2004, **116**(3):179-185.
4. Tinetti ME, Bogardus ST, Jr., Agostini JV: **Potential pitfalls of disease-specific guidelines for patients with multiple conditions.** *N Engl J Med* 2004, **351**(27):2870-2874.
5. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW: **Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance.** *JAMA* 2005, **294**(6):716-724.
6. May C, Montori VM, Mair FS: **We need minimally disruptive medicine.** *BMJ* 2009, **339**:b2803.
7. Eton DT, Ramalho de Oliveira D, Egginton JS, Ridgeway JL, Odell L, May CR, Montori VM: **Building a measurement framework of burden of treatment in complex patients with chronic conditions: a qualitative study.** *Patient Relat Outcome Meas* 2012, **3**:39-49.
8. Tran VT, Montori VM, Eton DT, Baruch D, Falissard B, Ravaud P: **Development and description of measurement properties of an instrument to assess Treatment Burden among patients with multiple chronic conditions.** *BMC Med* 2012, **10**(1):68.
9. Tran VT, Harrington M, Montori VM, Barnes C, Wicks P, Ravaud P: **Adaptation and validation of the Treatment Burden Questionnaire (TBQ) in English using an internet platform.** *BMC Med* 2014, **12**:109.
10. Vijan S, Hayward RA, Ronis DL, Hofer TP: **Brief report: the burden of diabetes therapy: implications for the design of effective patient-centered treatment regimens.** *J Gen Intern Med* 2005, **20**(5):479-482.
11. Gallacher K, May CR, Montori VM, Mair FS: **Understanding patients' experiences of treatment burden in chronic heart failure using normalization process theory.** *Ann Fam Med* 2011, **9**(3):235-243.

12. Gallacher K, Morrison D, Jani B, Macdonald S, May CR, Montori VM, Erwin PJ, Batty GD, Eton DT, Langhorne P *et al*: **Uncovering treatment burden as a key concept for stroke care: a systematic review of qualitative research.** *PLoS Med* 2013, **10**(6):e1001473.
13. Sav A, Kendall E, McMillan SS, Kelly F, Whitty JA, King MA, Wheeler AJ: **'You say treatment, I say hard work': treatment burden among people with chronic illness and their carers in Australia.** *Health Soc Care Community* 2013, **21**(6):665-674.
14. Denscombe M (ed.): **The Good Research Guide.** Buckingham: Open University Press; 1997.
15. Faugier J, Sargeant M: **Sampling hard to reach populations.** *J Adv Nurs* 1997, **26**(4):790-797.
16. Eton DT, Elraiyah TA, Yost KJ, Ridgeway JL, Johnson A, Egginton JS, Mullan RJ, Murad MH, Erwin PJ, Montori VM: **A systematic review of patient-reported measures of burden of treatment in three chronic diseases.** *Patient Relat Outcome Meas* 2013, **4**:7-20.
17. Eton DT, Oliveira DR, Egginton J, Mair FS, May C, Montori VM: **Understanding the burden of treatment in patients with multiple chronic conditions: Evidence from exploratory interviews.** *Qual Life Res* 2010, **19** (suppl 1)(131):ab. 1673.
18. Sav A, Kendall E, McMillan SS, Kelly F, Whitty JA, King MA, Wheeler AJ: **'You say treatment, I say hard work': treatment burden among people with chronic illness and their carers in Australia.** *Health Soc Care Community* 2013.
19. Sav A, King MA, Whitty JA, Kendall E, McMillan SS, Kelly F, Hunter B, Wheeler AJ: **Burden of treatment for chronic illness: a concept analysis and review of the literature.** *Health Expect* 2013.
20. Holsti OR: **"Content Analysis."** *The Handbook of Social Psychology.*: Reading MA: Addison-Wesley; 1968.
21. Lingard L, Albert M, Levinson W: **Grounded theory, mixed methods, and action research.** *BMJ* 2008, **337**:a567.
22. Reinert M: **Une méthode de classification descendante hiérarchique: Application à l'analyse lexicale par contexte.** *Cahiers de l'Analyse des Données* 1983, **3**:187-198.
23. Reinert M: **Classification descendante hiérarchique: un algorithme pour le traitement des tableaux logiques de grandes dimensions(1986)Un logiciel d'analyse lexicale [ALCESTE].** *Cahiers de l'Analyse des Données* 1985:471-484.
24. Noel-Jorand MC, Reinert M, Giudicelli S, Dassa D: **A new approach to discourse analysis in psychiatry, applied to a schizophrenic patient's speech.** *Schizophr Res* 1997, **25**(3):183-198.
25. Yang S, Kadouri A, Revah-Levy A, Mulvey EP, Falissard B: **Doing time: a qualitative study of long-term incarceration and the impact of mental illness.** *Int J Law Psychiatry* 2009, **32**(5):294-303.
26. Bradley EH, Curry LA, Devers KJ: **Qualitative data analysis for health services research: developing taxonomy, themes, and theory.** *Health Serv Res* 2007, **42**(4):1758-1772.
27. Harrison C, Britt H, Miller G, Henderson J: **Examining different measures of multimorbidity, using a large prospective cross-sectional study in Australian general practice.** *BMJ Open* 2014, **4**(7):e004694.
28. Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H: **A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology.** *Ann Fam Med* 2012, **10**(2):142-151.
29. Shippee ND, Shah ND, May CR, Mair FS, Montori VM: **Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice.** *J Clin Epidemiol* 2012, **65**(10):1041-1051.

30. Boyd CM, Wolff JL, Giovannetti E, Reider L, Weiss C, Xue QL, Leff B, Boulton C, Hughes T, Rand C: **Healthcare task difficulty among older adults with multimorbidity.** *Med Care* 2014, **52 Suppl 3**:S118-125.
31. Ridgeway JL, Beebe TJ, Chute CG, Eton DT, Hart LA, Frost MH, Jensen D, Montori VM, Smith JG, Smith SA *et al*: **A brief Patient-Reported Outcomes Quality of Life (PROQOL) instrument to improve patient care.** *PLoS Med* 2013, **10**(11):e1001548.
32. Russell LB, Suh DC, Safford MA: **Time requirements for diabetes self-management: too much for many?** *J Fam Pract* 2005, **54**(1):52-56.
33. Bohlen K, Scoville E, Shippee ND, May CR, Montori VM: **Overwhelmed patients: a videographic analysis of how patients with type 2 diabetes and clinicians articulate and address treatment burden during clinical encounters.** *Diabetes Care* 2012, **35**(1):47-49.
34. Malterud K: **Qualitative research: standards, challenges, and guidelines.** *Lancet* 2001, **358**(9280):483-488.
35. Duracinsky M, Herrmann S, Berzins B, Armstrong AR, Kohli R, Le Coeur S, Diouf A, Fournier I, Schechter M, Chassany O: **The development of PROQOL-HIV: an international instrument to assess the health-related quality of life of persons living with HIV/AIDS.** *J Acquir Immune Defic Syndr* 2012, **59**(5):498-505.
36. Britten N: **Qualitative interviews in medical research.** *BMJ* 1995, **311**(6999):251-253.
37. Ekman A, Litton JE: **New times, new needs; e-epidemiology.** *Eur J Epidemiol* 2007, **22**(5):285-292.
38. Flores G: **Language barriers to health care in the United States.** *N Engl J Med* 2006, **355**(3):229-231.
39. Woloshin S, Bickell NA, Schwartz LM, Gany F, Welch HG: **Language barriers in medicine in the United States.** *JAMA* 1995, **273**(9):724-728.
40. Streiner D, Norman G: **Biases in responding.** In: *Health Measurement Scales: a practical guide to their development and use.* Fourth Edition edn. Oxford: Oxford University Press; 2008: 103.
41. Guthrie B, Payne K, Alderson P, McMurdo ME, Mercer SW: **Adapting clinical guidelines to take account of multimorbidity.** *BMJ* 2012, **345**:e6341.
42. Montori VM, Brito JP, Murad MH: **The optimal practice of evidence-based medicine: incorporating patient preferences in practice guidelines.** *JAMA* 2013, **310**(23):2503-2504.

List of tables

Table 1: Characteristics of participants (n=1053)

Box 1: Healthcare tasks imposed on patients

Box 2: Factors that exacerbate the burden of treatment

Box 3: Consequences of healthcare tasks imposed on patients in their daily lives

List of figures

Figure 1: Flow of patients in the study

Figure 2: Taxonomy of the burden of treatment

List of Supplementary materials

Appendix 1: Patient associations, physician organizations and social media used for recruiting participants

Appendix 2: Questionnaire used on the website (English version)

Appendix 3: Extra tasks patient must perform because of their conditions (n=1053)

Appendix 4: Factors related to the nature and frequency of imposed tasks that exacerbate the burden of treatment (n=1053)

Appendix 5: Structural factors that exacerbate the burden of treatment (n=1053)

Appendix 6: Personal factors that exacerbate the burden of treatment (n=1053)

Appendix 7: Situational factors that exacerbate the burden of treatment (n=1053)

Appendix 8: Consequences of healthcare tasks imposed on patients in their daily lives (n=1053)

Appendix 9: Automatic textual analysis of English answers to open-ended questions (n=308).

Appendix 10a: Automatic textual analysis of answers to open-ended questions in French and Spanish.

Appendix 11: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of presence of multimorbidity (adjusted for age, gender, educational level).

Appendix 12: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of age (adjusted for presence of multimorbidity, gender, educational level).

Appendix 13: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of educational level (adjusted for presence of multimorbidity, gender, age).

Appendix 14: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of gender (adjusted for presence of multimorbidity, educational level, age).

Table 1: Characteristics of patients (n=1,053)

Characteristic	Total (n=1,053)	French- speaking patients (n=693)	English- speaking patients (n=315)	Spanish- speaking patients (n=45)
Age – yr	46 (14)	46 (14)	48 (14)	45 (12)
Female sex – no. (%)	768 (73)	474 (68)	262 (83)	30 (67)
Marital status – no. (%)				
Married or civil union	544 (52)	352 (51)	170 (54)	22 (49)
Live in partner	129 (12)	98 (14)	28 (9)	3 (6.7)
Single	252 (24)	166 (24)	70 (22)	16 (36)
Divorced	107 (10)	63 (9)	42 (14)	2 (4.4)
Widowed	21 (2)	14 (2)	5 (1.6)	2 (4.4)
Highest education level –no. (%)				
Primary/middle school	30 (2.8)	23 (3.3)	6 (1.9)	1 (2.2)
High school	305 (29)	249 (36)	44 (14)	12 (27)
College	718 (68)	421 (61)	265 (84)	32 (71)
Place where patients go most frequently for consultations – no. (%)				
Public hospital	465 (44)	378 (54)	65 (21)	22 (49)
Private hospital	83 (7.9)	61 (8.8)	17 (5.4)	5 (11)
General practitioner’s clinic	295 (28)	163 (24)	124 (39)	7 (16)
Specialist clinic	210 (20)	91 (13)	109 (35)	11 (24)
Presence of informal caregiver* – no. (%)	478 (45)	329 (47)	126 (40)	23 (51)
No. of medical appointments/month	6.2 (6.9)	7.1 (7.3)	4.4 (5.6)	4.9 (4.3)
No. of different physicians the patient regularly sees – no. (%)	3.3 (1.9)	3.3 (1.7)	3.5 (2.1)	3.2 (2)
No. of tablets/day	9.4 (9.3)	8.7 (8.4)	11 (11)	6.7 (7.3)
No. of injections/day	0.7 (1.3)	0.6 (1.4)	0.6 (1.2)	0.9 (1.7)
Time needed to organize medications per week – minutes	62 (137)	46 (95)	97 (202)	70 (70)
Time needed for self-monitoring per week – minutes	33 (96)	24 (60)	53 (160)	39 (56)
No. of conditions	2.4 (1.6)	2.2 (1.5)	2.7 (1.7)	2.3 (1.3)
Patients with >2 chronic conditions – no. (%)	662 (63)	409 (59)	224 (71)	29 (64)
Main chronic conditions**				
Diabetes	168 (16)	98 (14)	53 (17)	17 (38)
Other endocrine disorders	197 (19)	106 (15)	81 (26)	10 (22)
Lung diseases	142 (13)	52 (7.5)	84 (27)	6 (13)
High blood pressure or dyslipidemia	303 (29)	207 (30)	87 (28)	9 (20)
Heart diseases	125 (12)	57 (8.2)	64 (20)	4 (8.9)
Kidney or urological diseases	304 (29)	285 (41)	18 (5.7)	1 (2.2)
Gastrointestinal diseases	138 (13)	66 (9.5)	66 (21)	6 (13)
Stroke or cerebrovascular disease	35 (3.3)	20 (2.9)	15 (4.8)	0 (0)
Neurologic diseases	146 (14)	89 (13)	49 (16)	8 (18)
Rheumatologic disease	344 (33)	209 (30)	127 (40)	8 (18)
Cancer or malignant blood diseases	84 (8)	59 (8.5)	21 (6.7)	4 (8.9)
Psychiatric disease	128 (13)	56 (8.1)	60 (19)	12 (27)
Vision problems	79 (7.5)	46 (6.6)	29 (9.2)	4 (8.9)
Oto-rhino-laryngological problems	64 (6.1)	42 (6.1)	21 (6.7)	1 (2.2)
Skin diseases	89 (8.4)	58 (8.4)	21 (6.7)	10 (22)
Infectious disease	12 (1)	8 (1.1)	3 (0.9)	1 (2.2)
Systemic conditions	108 (10)	70 (10)	38 (12)	0 (0)
Other***	50 (4.7)	26 (3.7)	20 (6.3)	4 (8.9)

* Informal caregivers were family members or friends who helped the patient with healthcare tasks without payment for the care; ** A patient can have multiple chronic conditions. *** Other included non-malignant hematological conditions, thrombosis, obstetrical conditions, genetic disorders.

Box 1: Healthcare tasks imposed on patients

1.1. Management of medications

- Prepare and take drugs
- Plan and organize drug intake
- Follow specific precautions before, during or after medication intake
- Store medications at home
- Refill medication stock

1.2. Organizing and performing non-pharmacological treatment

- Access/use equipment
- Plan/perform physical therapy

1.3. Lifestyle changes

- Force myself to eat some foods
- Eliminate some foods
- Plan and prepare meals
- Be careful of ingredients in meals
- Organize physical exercise
- Perform some physical activities
- Give up some physical activities
- Change/organize sleep schedule
- Give up smoking
- Perform other lifestyle changes

1.4. Condition and treatment follow-up

- Plan and organize self-monitoring
- Plan and organize lab tests
- Precautions before/when performing tests
- Plan and organize doctor visits
- Remember questions to ask the doctor
- Organize transportation

1.5. Organize formal caregiver care

1.5. Paperwork tasks

- Take care of administrative paperwork
- Organize medical paperwork

1.6. Learning and developing an understanding of the illness and treatment

- Learning about my condition or treatment
- Learn to navigate the healthcare system

Box 2: Factors that exacerbate the burden of treatment

2.1. Nature, time required and frequency of healthcare tasks

2.2. Structural factors

- Access to resources
 - Pharmacy does not have the medication in stock
 - Access to lab test results
 - Access the right healthcare provider
 - Distance from healthcare facilities
 - Difficulty planning last-minute consultations
- No coordination between care providers
- Health center problems (e.g., wait times, parking near healthcare facilities)
- Not enough research done on my condition
- Insufficient or inadequate media coverage of my condition

2.3. Personal factors

- Beliefs
 - I'm anxious about performing tests and their results
 - I believe that some consultations are useless
 - I believe that some follow-up tasks are useless
 - I believe that my treatment is inefficient
 - I feel dependant on my treatment
 - My treatment conflicts with some of my religious beliefs
- Relationships with others (except healthcare providers)
 - I feel that I'm a burden for others
 - My loved ones overdo things/impose unnecessary precautions
 - My loved ones don't help me with my condition/treatment
 - I hide my condition or treatment from others
 - I have to regularly explain my conditions to others
 - Seeing other patients reminds me of what could happen to me in the future
- Relationships with healthcare providers
 - My physicians don't know about my condition/treatment
 - My physician doesn't take into account my context
 - Healthcare providers don't explain things to me
 - I feel that healthcare providers don't trust what I tell them
 - Healthcare providers don't take into account my psychological problems
 - Healthcare providers neglect some problems for others
 - For some healthcare providers, I'm just a condition and not a person

2.4. Situational factors

- Out of routine
 - Plan and organize travel
 - Store medications when not at home
 - Take medications when not at home
 - Access to structures or equipment when not at home
 - Pregnancy
- Other situational factors
 - Changing physicians
 - Organize my diet to accommodate other people
 - Follow my diet in the presence of other people

2.5. Financial factors

Box 3: Consequences of healthcare tasks imposed on patients in their daily lives

3.1. Lack of adherence

- Intentional non-adherence because of complexity
- Intentional non-adherence because of costs
- Non-intentional non-adherence and strategies not to forget treatment

3.2. Impact on professional, social, family life and leisure activities

- Opportunity cost in professional life. Coping with absence from work.
- My healthcare activities interfere with my career (e.g., I didn't get the job/promotion I wanted)
- Coping with judgment from others
- Treatment takes time/energy or requires precautions that interfere with family/friend commitments
- My healthcare activities interfere with my couple life
- Treatment takes time/energy or require precautions that interfere with leisure activities

3.3. Emotional impact

- Frustration because of not being able to do everything I want to
- Guilt associated with intentional non-adherence to treatment
- Treatment reminds me that I have a chronic condition

3.4. Financial impact of healthcare tasks imposed on patients

- Direct costs of treatment
- Indirect costs of treatment

Figure 1: Flow of participants in the study

* “Number of unique visitors” from May 22, 2013 to March 30, 2014, assessed by use of Google Analytics. “Unique visitors” can include study participants, non-eligible patients, physicians, researchers, members of patient associations, or curious visitors. Details on the method of calculating the number of users can be found at <https://support.google.com/analytics/answer/2992042?hl=en>

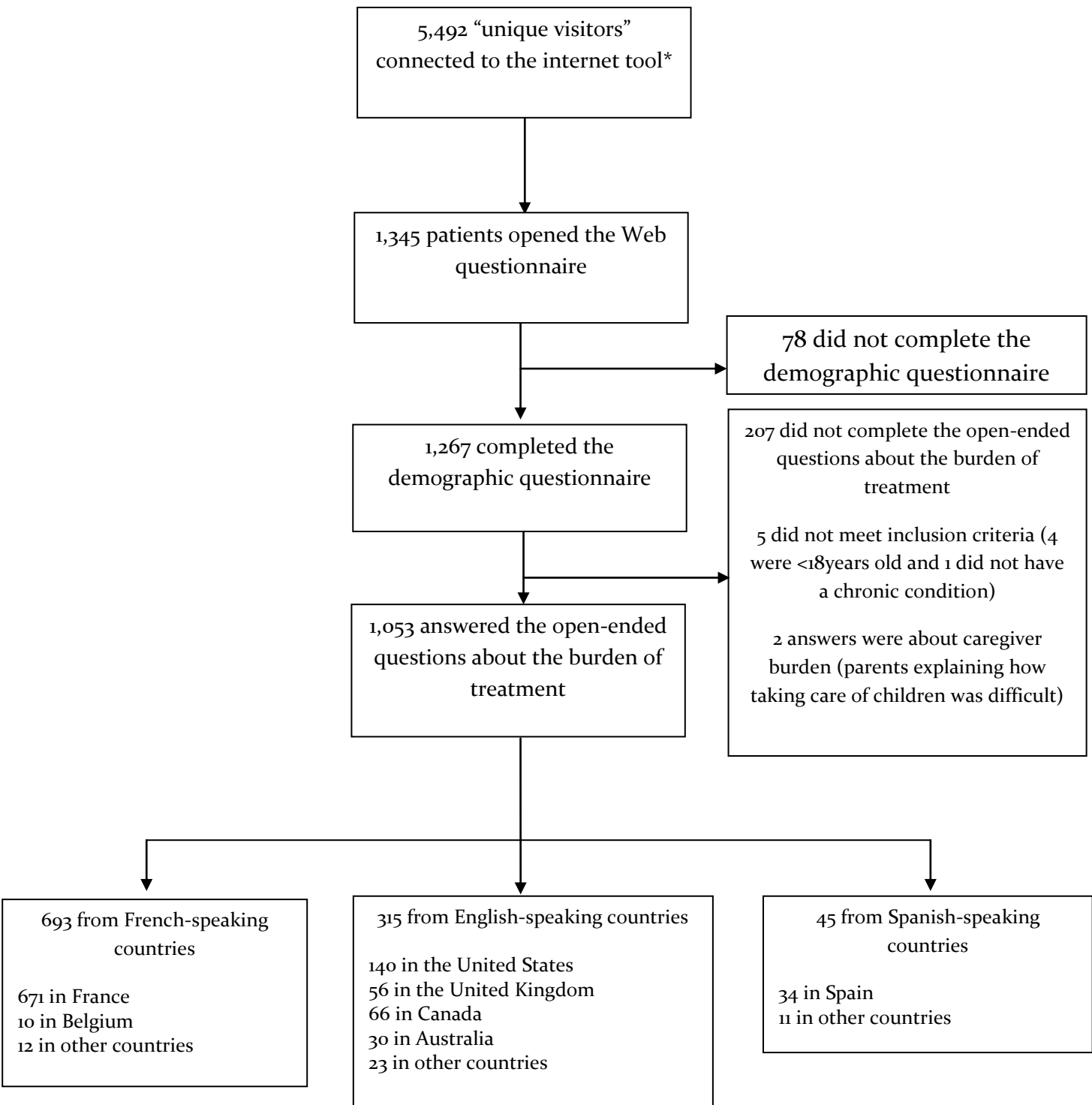
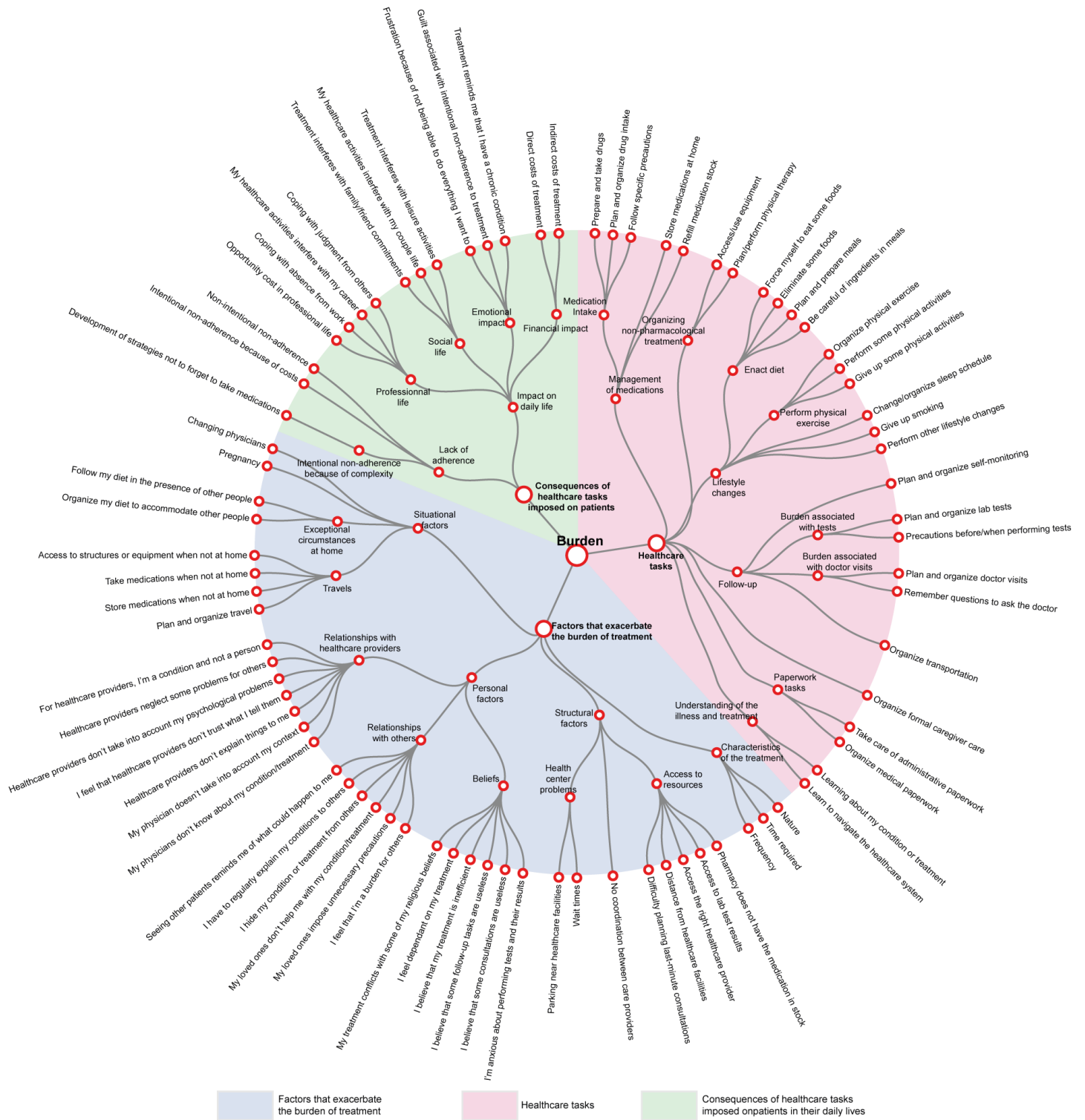


Figure 2: Taxonomy of the burden of treatment



Appendix 1: Patient associations, physician organizations and social media used for recruitment of participants

<p>Patient associations that agreed invite their members to participate in the study</p>	<p>ACS - Action contre les Spondylarthropathies AFLAR - Association Française de Lutte Anti-Rhumatismale AFS - Association France Spondylarthrite Alliance du cœur AMVF - Association des Malades des Vaisseaux du Foie ANDAR - Association Nationale de Défense contre l'Arthrite Rhumatoïde Associació Catalana de Fibrosi Quística Association des sclérodermiques de France Association Francophone pour Vaincre les Douleurs Association Patients du Canada Atrial fibrillation Association Australian Patient Association DEBRA International EMSP - European Multiple Sclerosis Platform Epilepsie France FNETH - Federación nacional de enfermos y trasplantados hepáticos Heart Sisters HTAP France Ligue Contre le Cancer Lupus Europe Lupus France Pain Alliance Europe Parenteral Nutrition-Down Under PHA Europe Renaloo SOS hépatite Wired4Life</p>
<p>Physician organizations that distributed information via e-mail about the study</p>	<p>Orphanet via their newsletter Cochrane collaboration via the Consumer Network</p>
<p>Social media</p>	<p>Twitter https://twitter.com/BurdenTreatment Mayo clinic Social Media Health Network http://network.socialmedia.mayoclinic.org/discussion/helping-patients-bear-the-burden-of-treatment Facebook page https://www.facebook.com/pages/Burden-of-Treatment/366884696759186</p>

Appendix 3: Extra tasks patient must perform because of their conditions (n=1,053)

Burden of treatment category	Example	Patients mentioning this burden In total - No. (%)	Patients mentioning this burden spontaneously - No. (%)*
Managing medications			
Prepare and take drugs correctly	“That syringe is difficult to handle - if you want to take off the cap, then the syringe sometimes unscrews from the syringe and you have spoiled a 200 EUR treatment”	237 (22)	123 (12)
Plan day accordingly to drug intakes	“Everything has to be planned for, you can no longer be spontaneous and have to make sure you take your pills at the proper times or you end up suffering the consequences.”	185 (18)	108 (10)
Following specific treatment precautions	“One medication requires you do not eat for two hours before taking the drug or one hour after. I take this twice a day. Fitting this into a schedule especially when travelling or visiting or having company is very difficult.”	291 (28)	127 (12)
Plan and organize refills	“This means I have to be on constant watch that I do not let supplies get too”	231 (22)	51 (4.8)
Store medications at home	“Medication storage is a bit inconvenient since they're supposed to be away from heat and light, and not in the kitchen or bathroom where there's moisture. They do take up a lot of space because there are many different bottles.”	111 (10)	20 (1.9)
Always have medication on person	“I have to ensure I have a spare Kalydeco tablet with me at all times in case I am delayed and have to eat a meal away from home, as this has to be taken with a fatty meal.”	111 (16)	53 (5.0)
Organizing and performing non-pharmacological treatment			
Access/use equipment	“We have built a wheelchair friendly home so that I can hopefully still be independent in my future years.”	35 (3.3)	28 (2.6)
Plan/perform physical therapy	“They tell us to do physical therapy everyday; it's tiring physically and mentally. Stretch up every day: when? In the evening I'm tired because of my day and my condition”***	52 (4.9)	38 (3.6)
Condition or treatment follow-up			
Plan and organize self monitoring	“I should test my blood every day. I do not. I go through phases where I test every day or every other day. And then sometimes I don't for weeks. It takes too long when I am running late for work in the morning.”	10 (0.9)	1 (0.1)
Plan and organize lab tests	“It's a day off work (impact on me). At least I can book an appt for lab work but it's still over an hour there. Hospital tests are good-about 3 hours start to finish. I usually don't plan anything else that day and make sure I have meals in the fridge the day before as I already know I won't have the energy to prep a meal”	184 (17)	66 (6.3)
Precautions before/when performing tests	“I have to lie down when I get blood taken. It's worse when I am fasting, which is usually the case. The lab nurses get frustrated because they only have one room at the lab with a bed. But if I don't tell them I need to lie down and I start to feel faint during the blood draw, they get mad that I didn't tell them.”	36 (3.4)	8 (0.8)
Plan and organize doctor visits	“That means that we have to manage our lives around the appointments. There is no coordination of the appointment arrangements. It appears that the relevant systems do not talk with one another. This means that the patient remains responsible for managing arrangements and that places more stress on the overall situation.”	428 (41)	205 (19)
Remember questions to ask the doctor	“Trying to get a drs appointment can be difficult so I have to have prepared for the appointment with a list [of questions] of what is needed”	8 (0.8)	1 (0.1)
Organize transportation	“I have to book wheelchair taxis several hours in advance and can only book to get to the appointment because I don't know how long I will wait for the appointment after arriving. Then when I call a taxi at the end of the appointment I can have a long wait.”	53 (5)	11 (1.0)
Lifestyle changes			
Force myself to eat some foods	“Trying to eat a high calorie diet and maintain normal blood sugar levels is extremely difficult. (...) It is also difficult trying to eat when you don't always feel well.”	23 (2.2)	15 (1.4)
Cut on some foods	“I hate dieting, it depresses me, and I do not enjoy having a limited diet. It makes me feel punished. I don't smoke, but I do drink wine most evenings, so being told to quit drinking is very frustrating, even though I know that is good for me.”	356 (34)	215 (20)

Plan and prepare meals	“I always have to plan ahead & cook my own meals. This takes considerable time & energy.”	80 (7.6)	54 (5.1)
Be careful of ingredients	“Extremely time consuming - its constant. I have to read every package for ingredients. I can't simply enjoy eating out with friends.”	65 (6.2)	39 (3.7)
Organize physical exercise	“I don't do it [following medical advice to practice regular physical exercise]. I work full time and have two children under three. (...) The medical advice to exercise and eat only the correct things and don't work too hard and don't get stressed is pathetically unrealistic.”	43 (4.1)	38 (3.6)
Perform some physical activities	“[Doctors tell me] I should swim but I hate water !”**	199 (19)	70 (6.6)
Give up some physical activities	“I don't follow their advice [to give up some activities]. If I did I would not get out of the bed.”	48 (4.6)	16 (1.5)
Change/organize my sleep schedule	“Making sure I get at least 9 hours sleep a day, which may mean canceling plans or failing to work if I had a night of insomnia.”	17 (1.6)	14 (1.3)
Give up smoking	“I stopped smoking before transplantation. It's been ten years and I still miss it”**	37 (3.5)	22 (2.1)
Perform other lifestyle changes	“I must avoid sunlight even though I need to walk daily for low impact exercise.”	110 (10)	80 (7.6)
Organize formal caregiver care			
Organize formal caregiver care	“But they [the nurses] kept coming to my home at eight in the morning, avoiding me to have any kind of love relationship and resulted into daily stress” **	8 (0.8)	6 (0.6)
Paperwork tasks			
Take care of administrative paperwork	“This part of it drives me insane with frustration. I tend to ignore bills, paperwork and admin when I'm ill and end up in trouble with various faceless organizations screaming at me on the phone for late payments of incomplete paperwork. pah!”	277 (26)	25 (2.4)
Organize medical paperwork	“I keep all my medical tests/reports/radiology/referrals in a file - it's very time consuming recording this, but vital so I can follow any changes in test results. I can then see if anything is changing & needs to be actioned. I keep a spreadsheet & data chart.”	51 (4.8)	19 (1.8)
Learning about the condition/treatment			
Learning about my condition or treatment	“I have tried to learn about my health conditions and when available have taken courses on the condition. I read about the conditions and follow the doctors recommendations regarding tests and medication”	79 (7.5)	47 (4.4)
Learn to navigate the healthcare system	“It's very complex. you have to be experienced with the whole medical system and then know how to advocate for yourself and navigate through a very complicated process”	64 (6.1)	17 (1.6)

*Spontaneously refers to patients mentioning the burden in the first broad open ended-question of the survey, prior to probes.

**Translated from another language

Appendix 4: Factors related to the nature and frequency of imposed tasks that exacerbate the burden of treatment (n=1,053)

Burden of treatment category	Example	Patients mentioning this burden In total - No. (%)	Patients mentioning this burden spontaneously - No. (%)*
Related to the nature of the tasks			
Treatment characteristics (taste, shape, size of tablets, discomfort of injections)	“I sometimes struggle to swallow my omeprazole (the biggest pill I have to take) and I can only take certain brands acetaminophen (...) because some are too hard to swallow and my domperidone tastes really bad.”	331 (31)	31 (2.9)
Getting the right dose is difficult (addition of tablets or fraction of tablets)	“I have to cut my tablets into quarters and add different dosages of the same medication to get the right amount”**	24 (2.3)	9 (0.9)
Appearance of medications can be confusing	“Different makes of the same med can look different which with brain-fog can be confusing”	23 (2.2)	4 (0.4)
Treatment is for whole life	“I will have to take medication for the rest of my life, there aren’t holidays for treatment”**	113 (11)	88 (8.3)
Medication side effects (real or imaginary)	“(…) two screws inserted during patella tendon realignment. The screws went right through the bone in to the calf muscles.”	323 (31)	174 (16)
Discomfort associated with tests	“Having blood drawn is often difficult. The vein in my arm often "hides", and can't be found at all, and blood has to be drawn from the back of my hand, which is very painful”	109 (10)	23 (2.2)
Discomfort associated with self-monitoring	“I should test my blood every day. I do not (...). It hurts.”	46 (4.4)	5 (0.5)
Discomfort associated with consultations	“The biggest problem I have is more the dignity of tests. When stripped for something like a heart echo, and other doctors walk into the room or you're not given a chance to cover yourself”	12 (1.1)	3 (0.3)
Time required to perform tasks			
Time needed to take/organize treatment	“Basic everyday treatments can sometimes be very time consuming. On a working day it would mean I have to get up extra early to complete everything before work”	97 (9.2)	68 (6.4)
Time needed for tests	“I am required to get monthly as well as 3-monthly blood tests. This requires organization and time.”	160 (15)	23 (2.2)
Time needed for self monitoring	“Wearing a heart rate monitor and logging activities and heart rate over a full day is time consuming and annoying. Writing that down consistently takes WAY more time than I would have imagined.”	50 (4.7)	10 (0.9)
Time needed for refills	“Repeated visits to pharmacists to make sure I have the correct meds take a considerable amount of time, especially as (...) pharmacists take considerable time to either dispense the prescription or have long queues to actually collect the medication.”	96 (9.1)	16 (1.5)
Time needed for doctor visits	“For many appointments, you must leave time for: getting to the appointment, finding parking, waiting for the appointment, seeing the doctor, getting back home. That can easily wipe out a morning or an afternoon”	206 (19)	48 (4.5)
Related to the number/frequency of the tasks			
Number/frequency of drug intakes	“Medications multiple times a day, and supplements multiple times a day, all scheduled so as not to interfere with each other and to meet different requirements”	298 (28)	239 (23)
Number/frequency of refills	“I go to the pharmacy so often that they know me by name! Not a good thing.”	125 (12)	39 (3.7)
Number/frequency of tests	“For annual review, it would be much better if all investigations could be ordered for the same day, then one visit and it is all done. But the hospital won't do that, so I have to make maybe 4 or 5 trips to different hospitals every year for this.”	270 (26)	224 (21)
Number/frequency of self monitoring acts	“Some days, it becomes a matter of why am I doing this every day. No one will look at it. Why do they want it? Why am I spending so much time trying to be thorough when nothing will change?”	48 (4.5)	33 (3.1)
Number/frequency of doctor visits	“[My condition/treatment] require a lot of medical follow ups, some every 4 months, some every year. Over the year, I average 1 to 2 appointments a month.”	274 (26)	226 (21)

*Spontaneously refers to patients mentioning the burden in the first broad open ended-question of the survey, prior to probes.

**Translated from another language

Appendix 5: Structural factors that exacerbate the burden of treatment (n=1,053)

Burden of treatment category	Example	Patients mentioning this burden In total - No. (%)	Patients mentioning this burden spontaneously - No. (%)*
Access to resources			
Pharmacy doesn't have my medication in stock	"It takes such a time to replace them-if for some reason the pharmacist does not have them in stock this can be extremely stressful."	37 (3.5)	6 (0.6)
Access to lab test results	"Doing tests is not fun. But it gets painful when results are not commented. Information is filtering through in dribs and drabs from doctors."**	18 (1.7)	5 (0.5)
Access the right healthcare provider	"Finding a family doctor that can and will take on a patient with multiple comorbidities is a major problem. Then finding one that is "smart" and able to cope with your conditions is even harder."	120 (11)	41 (3.9)
Distance from healthcare facilities	"I am able to drive but the distance and traffic factor into my comfort level for sure."	286 (27)	91 (8.6)
Difficulty planning last minute consultations	"Most doctors schedule a few weeks to a couple of months in advance, so it is hard to manage when you might NEED to see your doctor ahead of time."	28 (2.7)	10 (0.9)
No coordination between care providers			
No coordination between care providers	"That would be great if I could get any 2 doctors or specialist to agree on an activity that doesn't risk permanent ongoing damage"	100 (9.5)	39 (3.7)
Healthcare center problems			
Wait times	"Doctors appointments (...) are sometimes stressful due to the wait in the surgery [doctors' offices] when appointments are running late - as I run my own business I am also a busy person and regret the time wasted"	267 (25)	36 (3.4)
Parking near healthcare facilities	"The car parking is so expensive at the hospital I am forced to park illegally."	101 (9.6)	10 (0.9)
Research			
There is not enough research on my condition	"There is not enough research being done (...). There are no medications or devices approved for treating it. There are no accepted clinical guidelines for treatment. Everything is a stab in the dark. The biggest burden for me"	12 (1.1)	3 (0.3)
Media coverage			
Insufficient or inadequate media coverage of my condition	"My biggest problems at the moment are the media in my country, who are putting pressure on disabled people by characterizing us as "scroungers" who are really all fit to work. This stigma is difficult to bear."	4 (0.4)	0 (0.0)

Appendix 6: Personal factors that exacerbate the burden of treatment (n=1,053)

Burden of treatment category	Example	Patients mentioning this burden In total - No. (%)	Patients mentioning this burden spontaneously - No. (%)*
Beliefs			
I'm anxious about performing tests and their results	"It's a real stress that repeats every six months (...). A month before you think about it, during it it's terrible, and afterwards you never your results and even more complementary tests."	126 (12)	16 (1.5)
I believe that some consultations are useless	"I think it's ridiculous that I have to do follow-ups every month when my doctor doesn't appear to know much about my disease and definitely doesn't listen to me when I'm telling him how I'm doing and doesn't answer questions"	20 (1.9)	4 (0.4)
I believe that some follow-up tasks are useless	"You are asked to keep track of all of this, and then when you bring it in to your appointment, the doctor barely glances at it."	60 (5.7)	11 (1)
I believe that my treatment is inefficient	"Treatment is a burden, but worse: it doesn't work"	46 (4.4)	34 (3.2)
I feel dependant on my treatment	"Difficulty is to accept that, although I'm young, I'm dependant on medications to live." **	39 (3.7)	14 (1.3)
My treatment conflicts with some of my religious beliefs	"I can't follow all of my religious tasks, especially fasting as my urea gets higher when I don't drink" **	2 (0.2)	0 (0)
Relationships with others (except: healthcare providers)			
I feel that I'm a burden for others	"Having to rely on family for help is very difficult and makes me feel like a burden" / "I feel inadequate and that I'm not doing my share"	158 (15)	33 (3.1)
My loved ones overdo things/ impose unnecessary precautions	"Having family members tell me - stop doing this, or you'll get tired - that bothers me. I think I know when I'm tired. They don't realize that they make it worse by doing this (...) I want to be able to enjoy the "up" days, without constantly being told "slow down, you need to rest""	28 (2.6)	5 (0.5)
My loved ones don't help me with my condition/treatment	"In my couple, I live my condition alone. My husband doesn't care at all. Sometimes, I find that hard." **	15 (1.4)	5 (0.5)
I hide my condition or treatment from others	"Cystic fibrosis makes you sweat more and so I feel I have to always wear tops to try and hide this."	138 (13)	21 (2)
I have to regularly explain my conditions to others	"Can appear to be anti social. Having to explain why you can't eat or drink certain things can be tiring."	55 (5.2)	11 (1)
Seeing other patients reminds me of what could happen to me in the future	"Even then, sometimes I will get stuck in line behind someone getting blood work for a transplant - I feel glad I don't need a transplant, but really, that shouldn't be handled at the walk in lab."	6 (0.6)	2 (0.2)
Relationships with healthcare providers			
My physicians don't know about my condition/treatment	"One neurologist even insisted that pills couldn't help me (...), and that exercise would cure me. The level of ignorance is quite astounding (...). I find it very frustrating to keep getting it, since they shouldn't be giving advice about an illness which they do not understand at all."	82 (7.8)	41 (3.9)
My physician doesn't take into account my context	"Teach doctors that one way is NOT good for everyone. Individual care is vital as no two people are the same."	46 (4.4)	20 (1.9)
Healthcare providers don't explain things to me	"All is about explaining. If doctors took time to explain why they prescribed this medication, maybe it would be less a burden" (translated)	71 (6.7)	25 (2.4)
I feel that healthcare providers don't trust what I tell them	"It took me three years to convince a single medical professional that something was actually wrong with my knee, before I was finally told that I had arthritis. During that time I struggled with depression and worse, because I was so rudely dismissed."	44 (4.2)	16 (1.5)
Healthcare providers don't take into account my psychological problems	"It's like my physical health is more important that my mental health. They don't care that I'm in a bad mood as long as they can save my life. But what is my life worth if I'm unhappy. »	29 (2.7)	7 (0.7)
Healthcare providers neglect some problems for others	"I see a number of specialists who have knowledge of one or two organ systems and are very good at managing diseases affecting those systems, but fail miserably to effectively manage other problems. As a result, I almost don't take care of my other health problems."**	6 (0.6)	3 (0.3)
"For some healthcare providers, I'm just a condition and not a person"	"We as patients with chronic conditions are sometimes treated poorly by doctors and other healthcare members as drug seekers (...). I was horrified by the way I've been treated."	54 (5.1)	11 (1)

Appendix 7: Situational factors that exacerbate the burden of treatment (n=1053)

Burden of treatment category	Example	Patients mentioning this burden In total - No (%)	Patients mentioning this burden spontaneously - No (%)*
Out of routine			
Plan and organize travels	“Remember to take enough for total time I am going to be away when traveling. Get pharmacist to dispense extra. Take original pharmacy bottles when crossing borders”	198 (19)	66 (6.3)
Store medications when not at home	“If I fly abroad, all my hand baggage is taken up by my prescription supplies. This can be a problem if going for a two week holiday due to the restriction on fluid volumes that can be carried into the cabin. It also means I can't take any other essential hand baggage items”	134 (13)	21 (2)
Take medications when not at home	“The hardest thing is dealing with medications when changing time zones” “Remembering to take the pills when out (...) because I don't have the same memory triggers when I'm outside of my normal daily routine”	132 (12)	23 (2.2)
Access to structures or equipment when not at home	“Also getting access to a fridge is a difficulty (...) it effectively means we can't go camping any more. Some hotels are not good at allowing access to a fridge for medication. It is not nice to have to ask for medication in reception”	51 (4.8)	23 (2.2)
Pregnancy	“Let me tell you about pregnancy and type 2 diabetes! It is comical. Doctors appointments every two weeks between the gynecologist and the endocrinologist, 6 blood sugar readings and 4 insulin injections per day, the hematomas you can get when you have the bad luck of getting your vein pierced through”**	7 (0.7)	4 (0.4)
Other situational factors			
Changing physicians	“I also used to see an orthopedist, because (...) they can prescribe physiotherapy and I respond well to that as a treatment option, but unfortunately my doctor left the practice and her replacement isn't comfortable with chronic pain management.”	9 (0.8)	5 (0.5)
Organize my diet in function of other people	“You have to tell you hosts about your specific diet which is also a burden to them. When you eat with other people, you always have to choose a restaurant when you can also eat. It is not nice to impose such things to other people.”**	15 (1.4)	6 (0.6)
Follow my diet in presence of other people	“I cannot share much food with others, have to bring my own food, must deal with the hassle of having to explain to others why I cannot eat whatever I want, have the stigma of feeling weird and different from other people”	128 (12)	45 (4.3)

*Spontaneously refers to patients mentioning the burden in the first broad open ended-question of the survey, prior to probes.

**Translated from another language

Appendix 8: Consequences of healthcare tasks imposed to patients on their daily lives (n=1053)

Burden of treatment category	Example	Patients mentioning this burden In total - No (%)	Patients mentioning this burden spontaneously - No (%)*
Lack of adherence			
Intentional non-adherence because of complexity	“I decided early on that I just couldn't handle the shots in my belly because my pants waistband would just irritate the injection site reaction and cause an endless, low-level misery...so I just stopped taking that shot, and settled on a 6-day-a-week schedule, giving myself a day "off".	38 (3.6)	12 (1.1)
Intentional non-adherence because of costs	“I cannot afford to go to doctor visits every month, or every two months to refill my pain medications. I have had to stop taking my thyroid medications because I have no insurance or money to pay.”	26 (2.5)	6 (0.6)
Non-intentional non-adherence and strategies not to forget	“Also dealing with remembering to change the fentanyl patch and what side it was last on etc is taxing and requires a calendar nearby for noting these changes”	393 (37)	172 (16)
Impact on professional, social, family life and leisure activities			
Opportunity cost in professional life. Coping with absence from work	“My hospital is very exhaustive with testing (...) All of these are within work hours so I always have to take time off work. In addition, all of the communication with services has to take place while I'm at work.”	266 (25)	144 (14)
My healthcare activities interfere with my career (e.g. I didn't get the job/promotion I wanted)	“I gave up my career in chemistry due to not being able to cope with the flare ups. It has had a devastating effect on my career and difficulty to find work in areas that I wish to.”	140 (13)	73 (6.9)
Coping with judgment from others	“[When I] use my blue parking badge, I often get accused of abusing the system just because I don't have a wheelchair”	217 (21)	47 (4.4)
Treatment takes time/energy or require precautions that interfere with family/friends commitments	“One doctor suggested my problems would be solved if I cycled an extra 5km every day and gave up my family time on Sundays to exercise instead (...) These suggestions seemed arbitrary and unhelpful, and I must admit I disregarded them.”	171 (16)	91 (8.6)
My healthcare activities interfere with my couple life	“It is always a bit embarrassing to go to the bathroom with your purse when you're out with your friends or on a date, it looks like "she's having her period or that girl only thinks about making herself up" whereas actually it's for an injection.”**	22 (2.1)	11 (1)
Treatment takes time/energy or require precautions that interfere with leisure activities	“I no longer (...) spend time participating in leisure or recreation. Some of the medical stuff fits into the time those things would take, some gets done instead of taking a lunch break at work, the rest just doesn't get done.”	117 (11)	48 (4.5)
Emotional impact			
Guilt associated with intentional non adherence to treatment	“Every day I stick needles into my stomach and think about all the things I could have done to avoid the disease and all the things I should do but don't to lessen its impact on my life.”	48 (4.5)	8 (0.8)
Treatment reminds me that I have a chronic condition	“Thinking of all things I can't do, for example driving a motorcycle, remind me that I'm sick”	225 (21)	48 (4.5)
Financial impact of healthcare tasks imposed to patients			
Direct costs of treatment	“The costs of special equipment and special food, the cost of exercise programs, PT, OT, special shoes, a wheelchair accessible home with high property taxes, medication costs that are not covered, and the transportation to various appointments all take a big chunk out of income.”	347 (33)	67 (6.3)
Indirect costs of treatment	“I lose income if I do not work. I have no sick days or vacation days. I am a sole proprietor of small business open 7 days a week so if I have to take time off work that means I lose money for that day.”	120 (11)	23 (2.2)

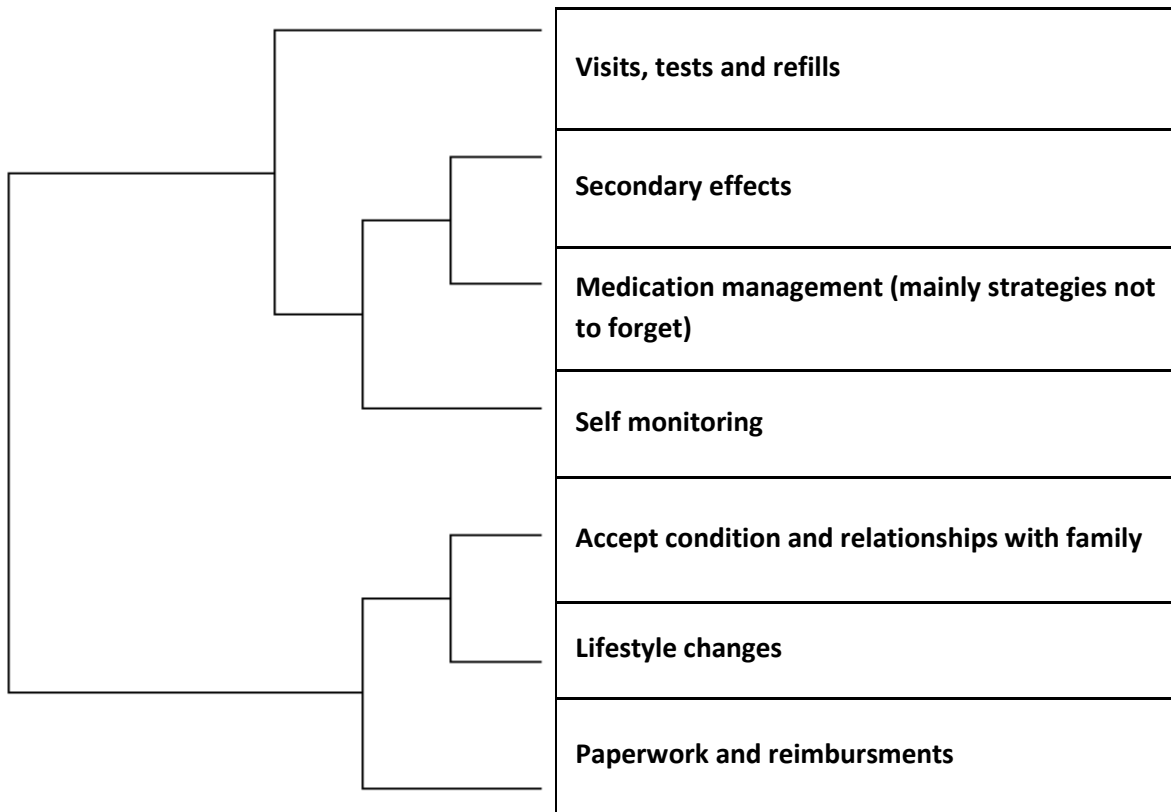
*Spontaneously refers to patients mentioning the burden in the first broad open ended-question of the survey, prior to probes.

**Translated from another language

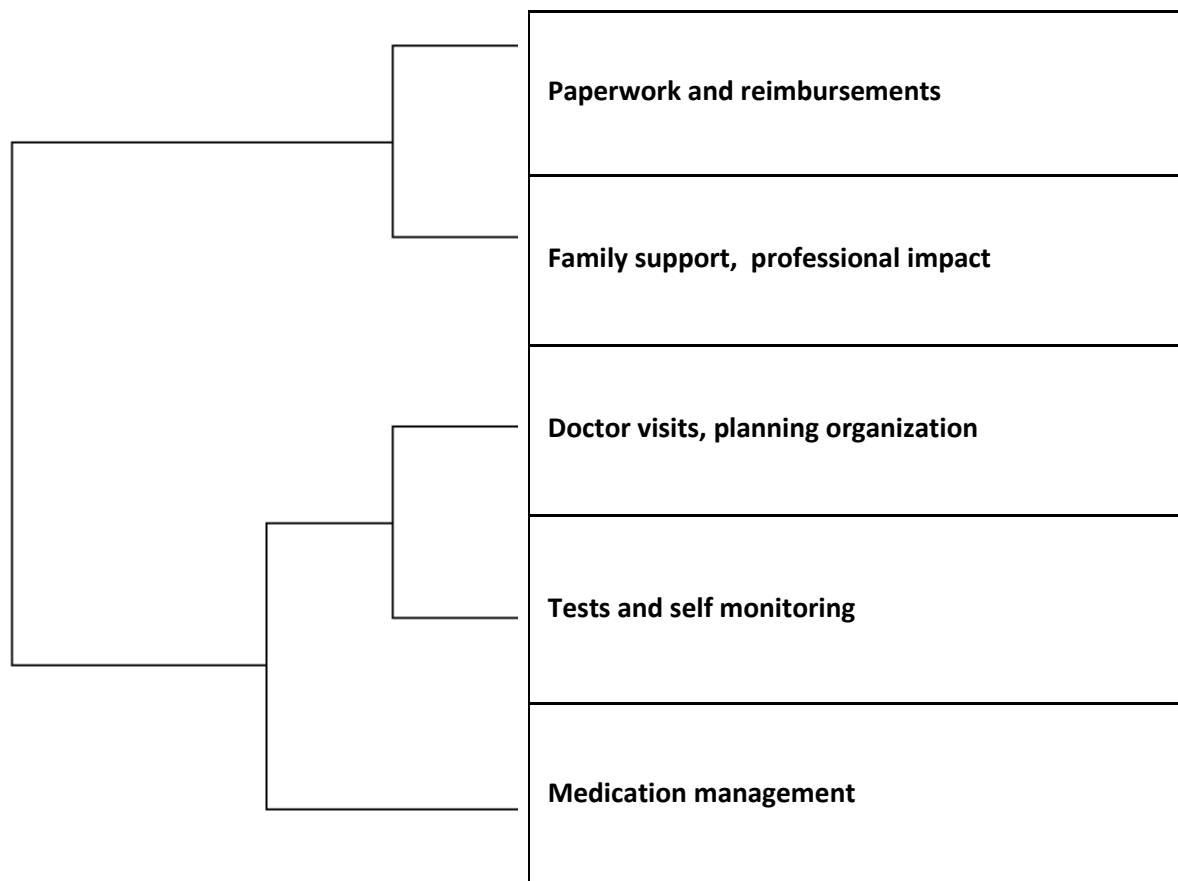
Appendix 9: Results of automatic textual analysis for English answers to open-ended questions (n=308). Words presented are words strongly associated with each class (chi-square>50) in descending order. Bold words are words most commonly used by participants. * Drug names were changed to international non-proprietary names

Tests and exams	Blood-test; wait; lab; visit; pressure; appointment; scan; blood work; time
Organize appointments; transportation	Appointment; come; wait; arrange; arrive; schedule; park
Refills	Pharmacy; refill; prescription; pick; call; order; mail; fax; email; request
Store ; organize during travels	Supply; bag; carry; fridge; bottle; mix; connect; travel; container; airport; pump; hotel; original; hook; plan; store; kitchen; infuse; luggage; nebulizer; oversea
Side effects, characteristics of drugs	Side-effect; injection; bruise; nausea; site; sting; painful; swallow; bee; inject; enoxaparin*; amitriptyline; cause; swell
Management of medications	Pill; stomach; breakfast; tablet; morning; remember; medication; taste; forget; empty; night; eat; alarm; dose; day
Lifestyle changes (diet and exercise)	Exercise; diet; food; alcohol; eat; weight; walk; smoke; rehab; calorie; appetite; healthy; gym; joint; arthritis; drink; cook
Relationships with others (family friend)	Family; member; old; feel; friend; life
Relationships with others (society)	Condition; chronic; understand; health; person; mental; treat; patient; management; fear; society
Paperwork	Disability; assistance; cover; insurance; benefit; medicare; apply; approve; patience; appeal; company; qualify; paperwork
Financial costs	Insurance; bill; pay; expense; cover; pocket; income; claim; government; cost; dollar; financial; coverage; thousand; park

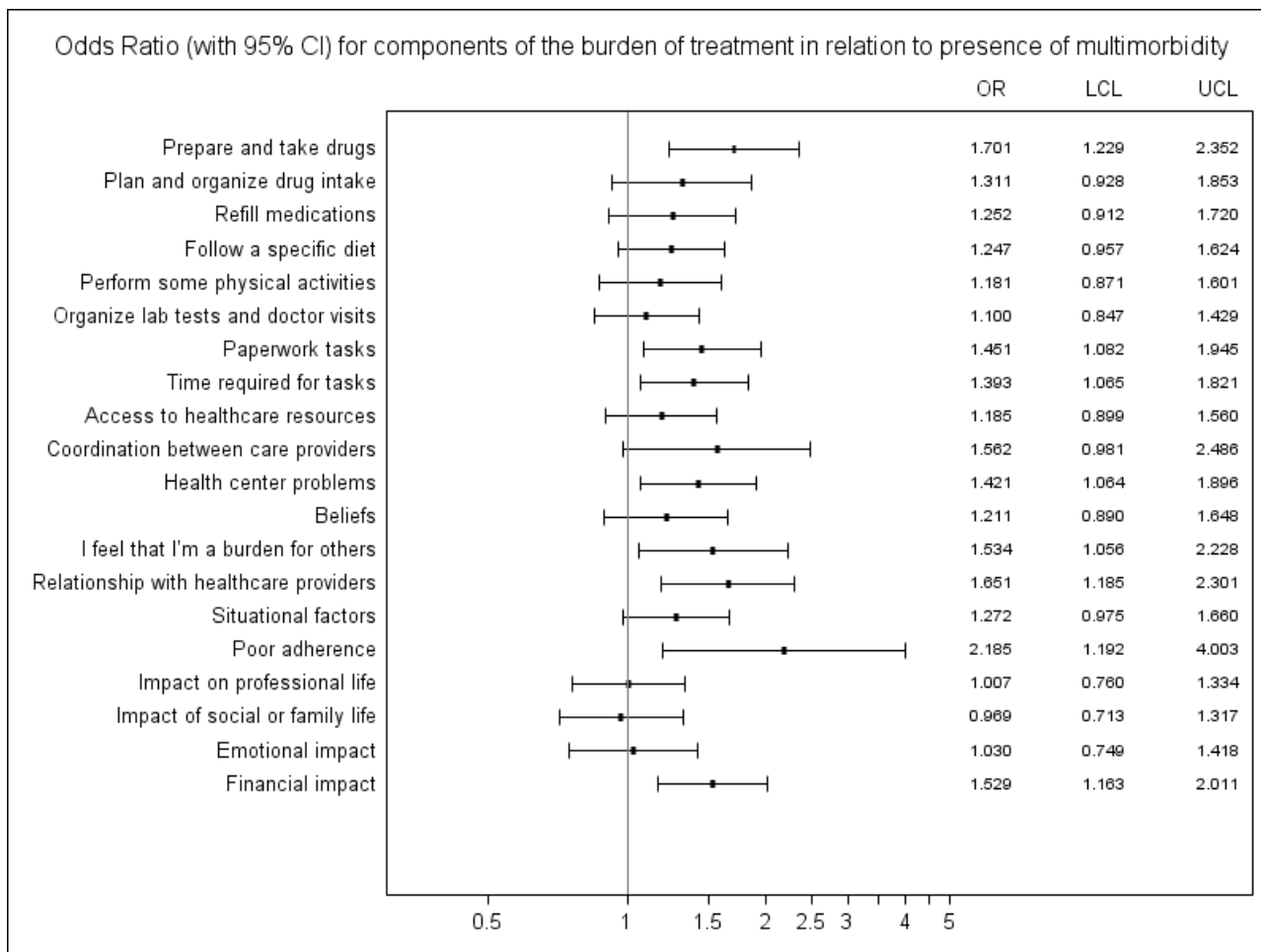
Appendix 10a: Results of automatic textual analysis for answers to open ended questions in French (n=690).



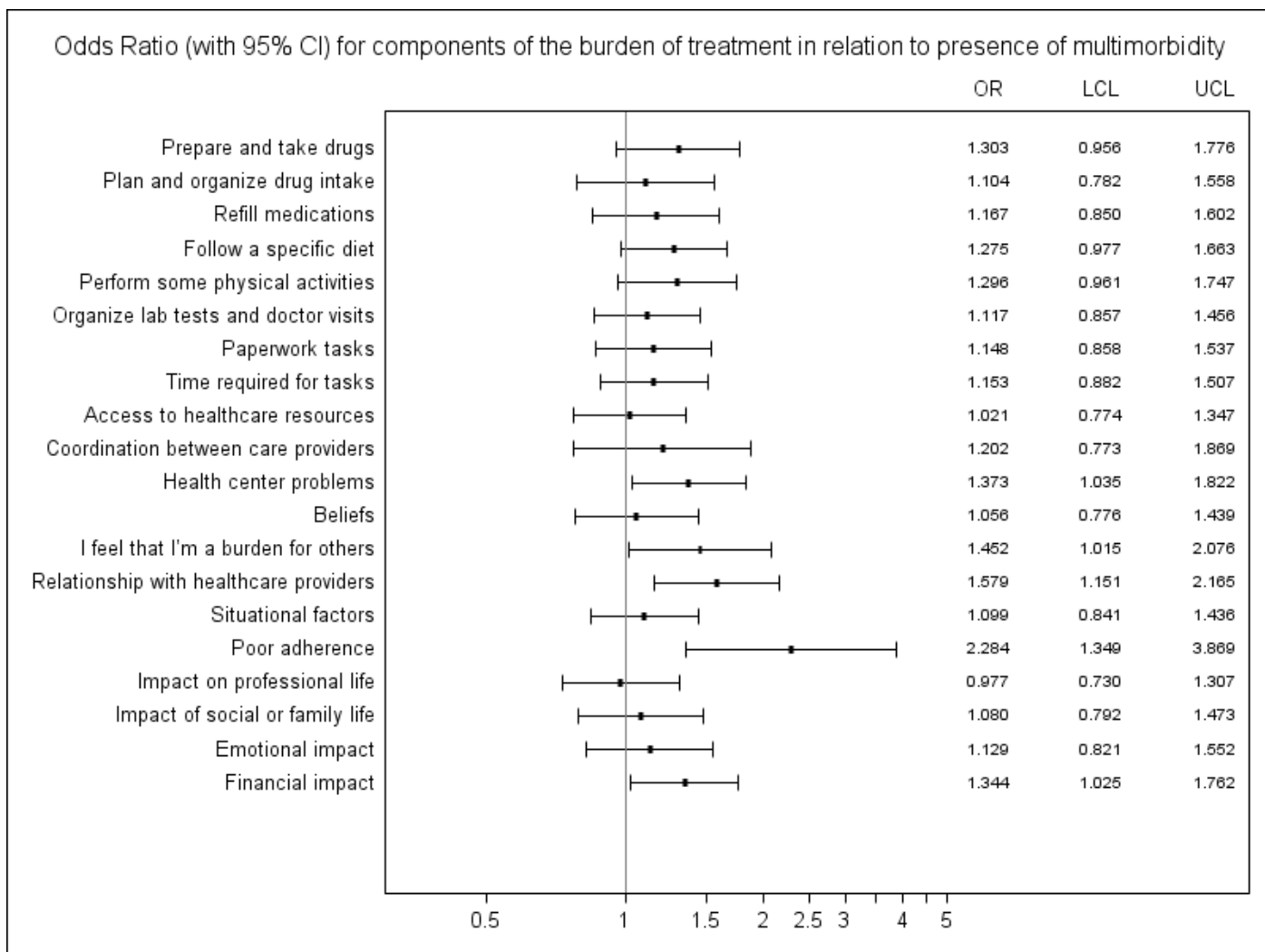
Appendix 10b: Results of automatic textual analysis for answers to open ended questions in Spanish (n=42).



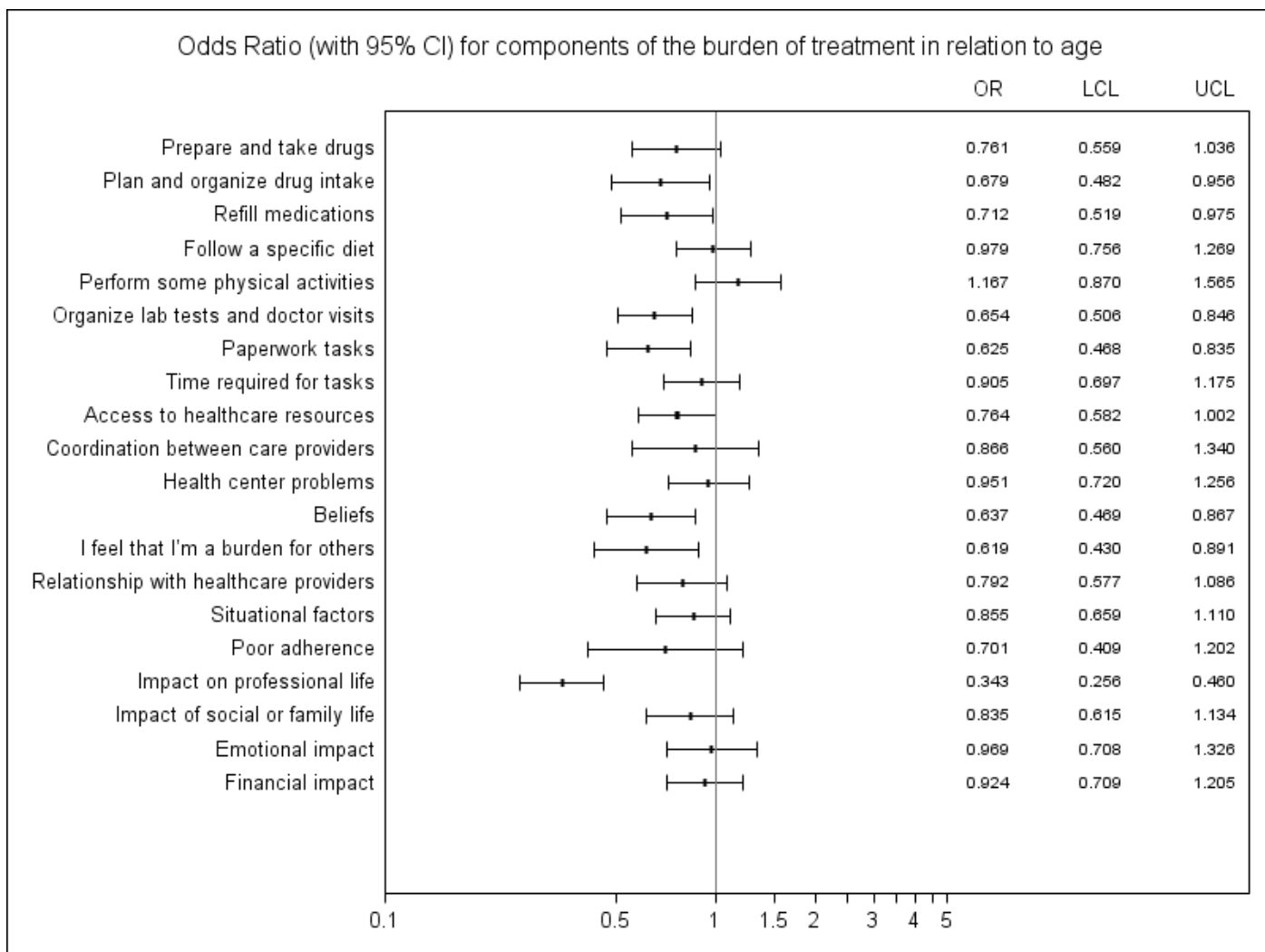
Appendix 11a: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of presence of multimorbidity (defined as patients with 2 or more chronic conditions) (adjusted for age, gender, educational level). Higher OR indicates that patients with >2 chronic conditions elicited the burden more often than those with one chronic condition.



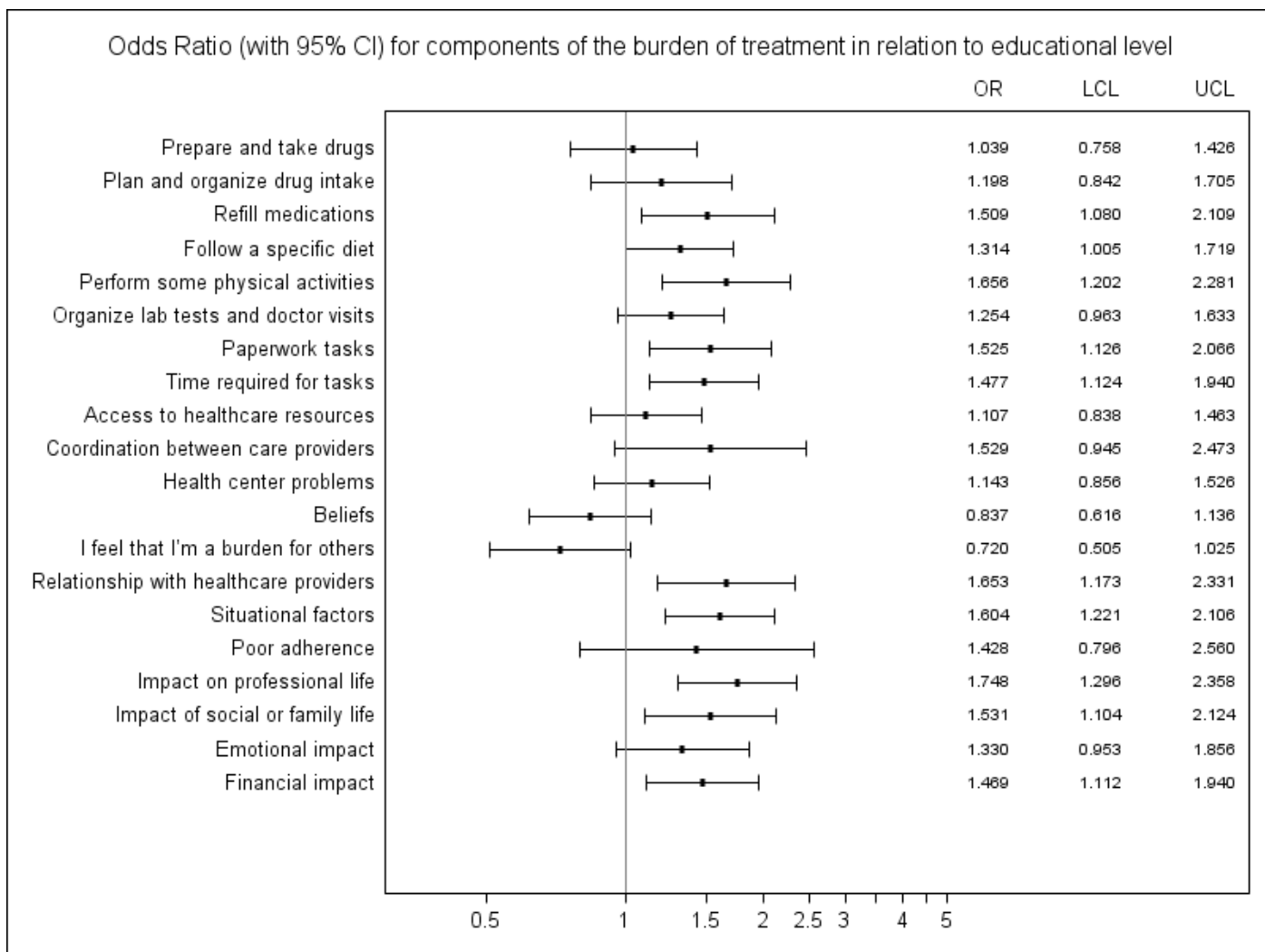
Appendix 11b: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of presence of multimorbidity (defined as patients with 3 or more chronic conditions) (adjusted for age, gender, educational level). Higher OR indicates that patients with >3 chronic conditions elicited the burden more often than those with one chronic condition.



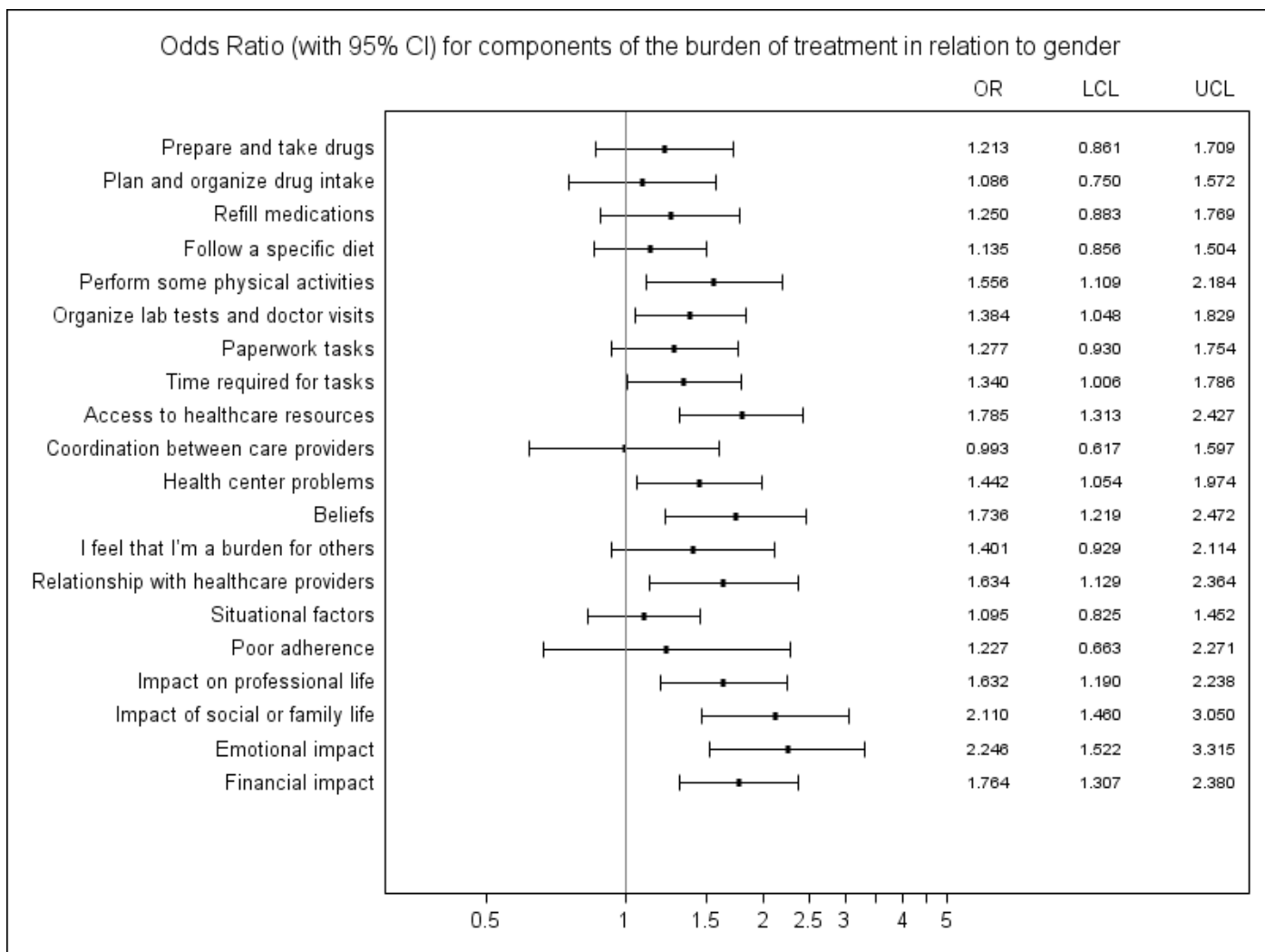
Appendix 12: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of age (adjusted for presence of multimorbidity, gender, educational level). Higher OR indicates that patients >50 years old elicited the burden more often than those < 50 years old.



Appendix 13: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of educational level (adjusted for presence of multimorbidity, gender, age). Higher OR indicates that patients with higher educational level elicited the burden more often than those with lower educational level.



Appendix 14: Odds ratios (with 95% CI) for components of the burden of treatment elicited by patients in terms of gender (adjusted for presence of multimorbidity, educational level, age). Higher OR indicates that female participants elicited the burden more often than male participants.



6. Taille d'échantillon dans les études qualitatives

Dans le chapitre précédent, nous avons décrit comment nous avons réalisé une étude qualitative pour explorer les composantes du fardeau du traitement. A notre connaissance, c'était l'une des premières études à utiliser un outil internet pour recueillir des données qualitatives. Lors de cette étude, nous avons recueilli les avis de nombreux patients venant de pays différents. Nous avons saisi l'opportunité de cette étude pour nous pencher sur le concept de saturation des données dans les études qualitatives.

6.1. Définition de la saturation des données

Le concept de saturation des données, a été défini par Glaser et Strauss en 1967, comme le fait que « l'analyse de nouvelles données ne développe pas davantage la description des thèmes et n'enrichit pas la théorie en cours d'étude » [96]. Selon, Glaser et Strauss, le point de saturation des données est déterminé par : 1) les limites empiriques des données ; 2) l'intégration et la richesse de la théorie développée ; et 3) la sensibilité de l'analyste. Depuis cette définition, le concept de saturation des données a été largement vulgarisé et représente, de nos jours, le principal critère d'arrêt du recueil des données dans les études qualitatives.

Dans la littérature, on ne trouve aucune méthode définie pour estimer a priori le point de saturation des données, et de fait, la taille d'échantillon nécessaire. Le recueil des données études ne doit être arrêté que lorsque le chercheur, après de fréquents aller retours entre recueil et analyse des données, a l'impression d'avoir « tout entendu »

[97]. Cependant, même si la plupart des recommandations en recherche qualitative insistent sur le fait que la taille de l'échantillon ne peut être définie que de manière inductive, la plupart des organismes de financement de la recherche ou des comités d'éthique demandent aux chercheurs de spécifier, a priori, le nombre de patients qu'ils comptent inclure [98, 99]. Dans ces cas-là, la taille de l'échantillon est généralement fixée en fonction d'études antérieures portant sur le même sujet [100], de recommandations empiriques, ou en fonction de l'expérience des investigateurs [99, 101].

6.2. Méthodes d'estimation du point de saturation des données

Bien que la saturation des données soit un concept simple à appréhender en pratique, l'estimation, avant toute recherche, du point de saturation des données est difficile. Dans la littérature, nous pouvons retrouver un nombre important de recommandations sur le nombre d'interviews ou de focus groups à réaliser, en fonction du type d'analyse (analyse fondée sur la théorie ancrée, phénoménologie, ethnographie, etc.). Par exemple, Morse et al. déclare qu'en général : 1) 6 participants permettent d'obtenir la saturation des données dans les études où l'analyse est phénoménologique ; 2) 35 participants permettent d'obtenir la saturation des données dans les études où l'analyse est fondée sur la théorie ancrée ou en ethnographie ; et que 3) 100 à 200 participants permettent d'obtenir la saturation des données dans les études éthologiques [102]. La plupart des autres auteurs recommandent généralement des tailles d'échantillon similaires, toujours fondées sur l'expérience des chercheurs [103].

Plusieurs études ont décrit le nombre d'interviews conduites dans les études qualitatives. En particulier, Mason et al. a décrit la taille d'échantillon d'études conduites dans le cadre de thèses d'université (PhD) soutenues au Royaume-Uni entre 1716 et 2009, toutes disciplines confondues [104]. L'auteur a identifié 2533 thèses ayant utilisé une méthode qualitative. Parmi ces thèses, 560 avaient précisé la taille de l'échantillon dans le résumé. La taille d'échantillon allait de 1 cas (histoire d'une vie) à 95 cas. La taille moyenne des études était de 31 participants (SD=18.7). Un résultat notable était la proportion plus importante d'échantillons dont la taille était un multiple de 10. Ceci est une preuve que la taille d'échantillon dans une recherche qualitative ne repose pas uniquement sur le seul concept de saturation des données.

Guest et al. est le premier auteur à avoir tenté de répondre à la question de la taille d'échantillon dans les études qualitatives [98] en étudiant le nombre cumulé de thèmes identifiés en fonction du nombre de patients. Les auteurs ont utilisé les données d'une étude impliquant 60 femmes vivant au Ghana et Nigeria parlant de relations sexuelles dont l'objectif était de déterminer les moyens adéquats d'aborder avec les patientes les méthodes pour prévenir la transmission du virus du VIH. Dans cette étude, les auteurs ont considéré comme point de saturation le point où « l'ajout de nouvelles données ne modifiait plus les thèmes identifiés ». Pour cela, les auteurs ont groupé les verbatim des participantes par 6. Ensuite, ils ont analysé les verbatim du premier groupe, formant un premier « livre de thèmes ». Puis ils ont analysé chaque groupe suivant, en notant si l'ajout de nouvelles informations modifiait ce « livre de thèmes ». Au final, les auteurs ont identifié 109 thèmes dans leur étude ; 80 (73%) avaient été identifiés après l'analyse du 1^{er} groupe de 6 interviews et 100 (92%) avaient été identifiés après l'ajout d'un second groupe de 6 interviews. La **figure 6** illustre les

résultats de leur étude. Les auteurs concluent que, dans le contexte de leur étude, 12 interviews permettaient d'identifier >90% des thèmes possibles.

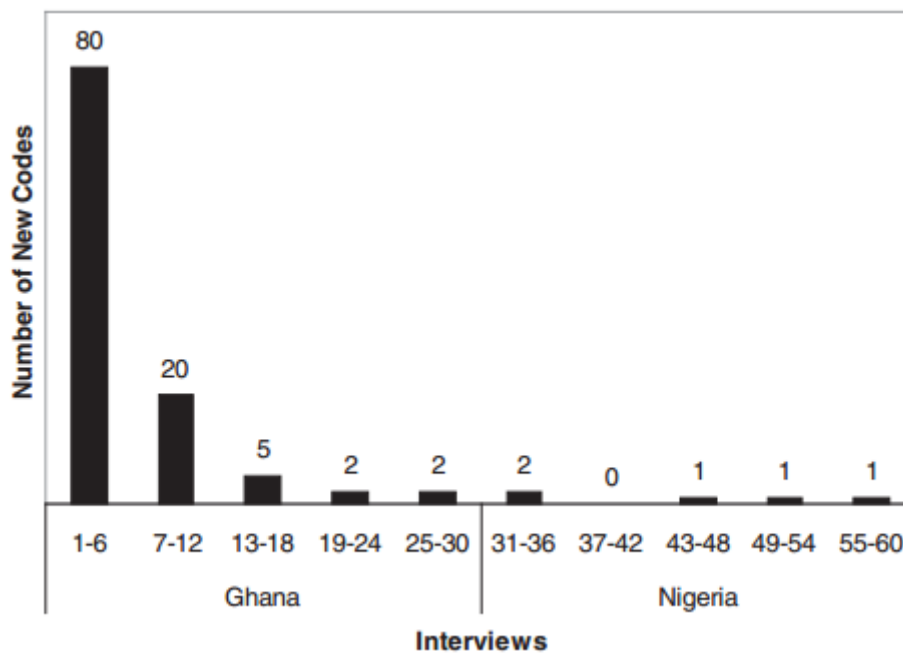


Figure 6 : Nombre d'interviews nécessaire pour obtenir la saturation des données, adapté de la source [98]

Francis et al. a abordé la question de l'estimation du point de saturation des données avec un autre angle. Les auteurs se sont intéressés à la définition pratique de ce point. En effet, que signifie en pratique « qu'aucun nouveau thème n'a été identifié » ? Si le second participant inclus dans l'étude est très similaire au premier et qu'il ne mentionne aucune nouvelle information, doit-on conclure que nous avons atteint la saturation des données après 2 interviews ? Pour pallier à ce problème, les auteurs ont revu la définition de la saturation des données comme : « l'absence de nouvelle information émergeant de l'analyse après X interviews ». Ainsi, les auteurs proposent une méthode pour s'assurer que la saturation des données a été établie reposant sur : 1) la définition a priori d'un échantillon initial d'au moins 10 participants ; 2) l'ajout

consécutif de participants à l'échantillon initial jusqu'à ce qu'aucune information nouvelle ne soit identifiée lors de 3 interviews consécutives. Pour tester leur méthode, les auteurs ont analysé les données d'une étude qualitative impliquant 17 patients ayant un proche atteint de maladie de Paget et cherchant à décrire leurs croyances, dans le but d'évaluer l'acceptabilité d'un dépistage génétique pour cette maladie. Au total, les interviews ont permis de décrire 35 croyances différentes. L'utilisation d'un critère d'arrêt a permis de ne pas arrêter le recrutement trop tôt, tout en ayant l'assurance d'avoir une stabilité dans le codage. Cependant, dans cette étude, les auteurs n'ont pas étudié l'impact d'une variation du critère d'arrêt sur le point de saturation des données.

6.3. Estimer le point de saturation des données dans les études qualitative en ligne

Etant donné que nous avons réalisé une des premières études qualitatives en ligne, il n'y avait ni recommandation ni étude antérieure pouvant nous aider à prédire le nombre approximatif de sujets à inclure. De part leur nature, les études en ligne présentent des spécificités par rapport aux études par interviews qui limitent la transposition des résultats des travaux précédemment : 1) le recrutement des patients est « instantané », permettant de recueillir un grand nombre de données provenant de participants géographiquement distants à faible cout ; et 2) le questionnaire rigide produit des données moins riches que lors d'entretiens face-à face avec le participant oblige a recruter un plus grand nombre de sujets pour obtenir une grande diversité de thèmes.

L'objectif de ce travail était donc d'évaluer l'impact de différentes stratégies d'échantillonnage (nombre de sujets à inclure et caractéristiques des sujets à inclure) sur la vitesse à laquelle nous arrivions à saturation des données dans une étude qualitative en ligne.

6.3.1. Méthode

Nous avons utilisé les données issues de notre étude qualitative internationale visant à décrire les différentes composantes du fardeau du traitement. Lors de cette étude, nous avons associé à chaque participant les thèmes qu'il avait mentionné au cours de l'étude. Ceci nous a permis de déterminer le nombre de thèmes mentionnés cumulés en fonction du nombre de patients inclus dans l'étude.

6.3.1.1. Choix du seuil de saturation des données

Guest et al. avait défini comme saturation des données le fait d'avoir identifié >90% des thèmes possibles indépendamment de leur nature ou fréquence [98]. De fait, nous avons choisi la même définition dans notre travail. Nous avons également défini la notion de « thème fréquent » : un thème fréquent est un thème qui a été mentionné par au moins 2.5% des patients dans l'étude expérimentale. Cette limite a été choisie afin de pouvoir relativiser l'impact de thèmes mentionnés par un très petit nombre de participants dans l'estimation du point de saturation des données.

6.3.1.2. Analyses statistiques

Nous avons réalisé des simulations de Monte Carlo afin de ré-échantillonner les données de notre étude expérimentale. En utilisant ces simulations, nous avons déterminé : 1) la probabilité d'identifier un nouveau thème ou de manquer un thème fréquent ; 2) l'efficacité de deux stratégies pour décider d'arrêter le recueil de données (choix d'une taille fixe a priori ; et stratégie décrite par Francis et al. [105] avec comme

critère d'arrêt 3 ou 10 interviews sans nouveau thème identifié); et 3) l'efficacité de deux stratégies pour décider quels patients à inclure dans l'étude (échantillonnage raisonné en fonction du pays de résidence des participants ou de leurs maladies et échantillonnage « aléatoire »).

Toutes les simulations ont été répétées 10 000 fois. Les participants ont été inclus sans remise car nous avons considéré que notre échantillon initial n'était pas représentatif d'une plus grande population. Les analyses ont été réalisées à l'aide du logiciel R version 3.1.

6.3.2. Résultats

Dans l'étude expérimentale, nous avons identifié 123 thèmes différents lors de l'analyse. Tous les thèmes avaient été mentionnés au moins 1 fois après le 681^{ème} patient. Le point de saturation des données, défini comme l'identification de 111 (90%) des thèmes avait été atteinte après le 192th patient. Parmi les 123 thèmes identifiés, 88 (71.5%) ont été définis comme « fréquents ». Tous les thèmes fréquents avaient été identifiés après le 149^{ème} participant.

Les probabilités médianes d'identifier au moins un nouveau thème de 32%, 24% et 12% pour le 30^{ème}, 50^{ème} et 100^{ème} sujet inclus. Comparativement. Les probabilités médianes de manquer un thème fréquent étaient de 17%, 6% et 1% pour le le 30^{ème}, 50^{ème} et 100^{ème} sujet inclus.

Dans l'étude de simulation, la saturation des données était obtenue après inclusion de 150 participants [IQR=91-93%]. Si l'on ne considérait que la première question ouverte, nous n'avons pas obtenu la saturation des données avec l'inclusion de 250 participants [IQR =83-86%]. Ces résultats sont illustrés par la **Figure 7**.

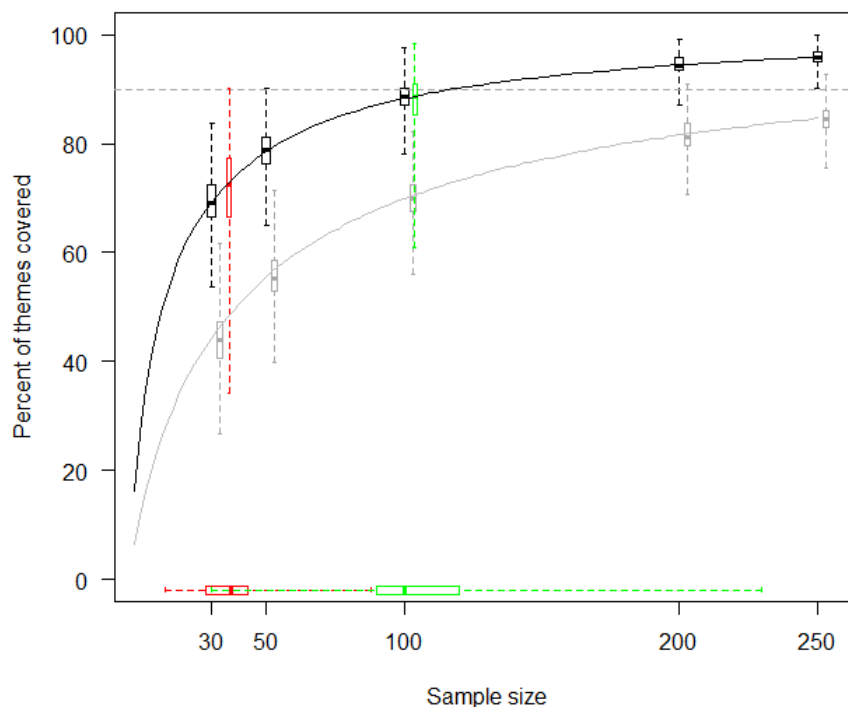


Figure 7 : Proportion de thèmes identifiés en fonction du nombre de participants. La courbe noire représente la proportion moyenne de thèmes identifiés avec un questionnaire composé d'une question générale et de multiples questions plus précises interrogeant sur des points importants identifiés dans la littérature. La courbe grise représente la proportion moyenne de thèmes identifiés avec un questionnaire composé d'une unique question générale. Les boîtes à moustache rouges et vertes représentent les échantillons recrutés à l'aide de la méthode de Francis [105] avec un critère d'arrêt de respectivement 3 et 10 participants n'évoquant pas de nouveau thème.

Nous avons évalué la stratégie proposée par Francis et al. pour estimer le point de saturation des données. Nous n'avons pas obtenu la saturation des données en utilisant comme critère d'arrêt le fait d'avoir 3 participants consécutifs ne mentionnant aucun nouveau thème : le critère d'arrêt était vérifié après inclusion d'un nombre médian de 37 [IQR=28-43] patients, mentionnant 72% [IQR=66%-77%] des thèmes. L'utilisation d'un critère d'arrêt de 10 patients ne mentionnant pas de nouveaux thèmes donnait de meilleurs résultats mais ne parvenait pas à atteindre la saturation des données (**Figure 7**).

Nous avons ensuite évalué si l'addition de patients étrangers à un échantillon initial de patients français permettait d'atteindre plus rapidement la saturation des données. Nous avons montré que recruter des patients étrangers augmentait significativement le nombre de thèmes mentionnés pour des tailles d'échantillon <100 patients. Cependant le gain était faible. Enfin, nous avons fait de même afin de savoir si l'addition de patients présentant les maladies les moins représentées dans l'échantillon initial permettait d'atteindre le point de saturation des données plus rapidement. Nous n'avons pas montré de différence significative concernant le nombre de thèmes mentionnés.

6.3.3. Discussion

La saturation des données est un concept facile à comprendre mais difficile à évaluer en pratique. Le point de saturation des données dépend du sujet d'étude, des participants, des méthodes de recueil et d'analyse des données [103]. Dans ce travail, nous avons ré-analysé les données d'une étude qualitative impliquant 1053 participants dans 34 pays différents. Nous avons montré que l'inclusion de 150 patients tirés aléatoirement permettait de s'assurer de la saturation des données. Des tailles d'échantillon plus réduites sont possibles mais exposent au risque de manquer des thèmes fréquents.

A notre connaissance, c'est la première étude à avoir utilisé des techniques de simulation pour déterminer une estimation du nombre de thèmes cumulés et du point de saturation des données. Cependant, cette étude a plusieurs limites. Tout d'abord, les estimations ont été obtenues à partir d'un unique jeu de données. Il est probable que les résultats auraient été différents dans d'autres contextes, avec des participants

différents, des méthodes de recueil et d'analyse de données différentes. Ensuite, nous avons utilisé une définition très arbitraire de la saturation des données, sans rapport avec la définition originale de Glaser et Strauss [96]. Enfin, lorsque nous recherchés la vitesse à laquelle le point de saturation des données était atteint en fonction des caractéristiques des patients inclus, nous avons été limités par les données disponibles ; un vrai échantillonnage raisonné repose sur l'inclusion de patients en fonction des thèmes déjà évoqués, pour les préciser [106].

La saturation des données est une notion élastique dépendant à la fois de l'objectif de la recherche et du niveau d'analyse désiré : de nouvelles données vont toujours apporter quelque chose de nouveau, apportant plus de grain aux thèmes identifiés. Idéalement, le recrutement pourrait être infini ; cependant la rentabilité de l'inclusion de patients supplémentaires (en termes de temps investigateur, temps patient, et coût de la recherche) diminue en fonction du nombre de patients déjà inclus. Notre étude offre un moyen pour des chercheurs réalisant des études qualitatives en ligne d'estimer la rentabilité d'inclure ou pas de nouveaux participants en fonction du nombre de participants déjà inclus.

6.3.4. Article

Tran VT, Porcher R, Falissard B, Ravaud P. Determining sample size in Web-based qualitative studies: the example of a multinational study about the burden of treatment. *Soumis au Journal of Clinical Epidemiology*

Determining sample size in Web-based qualitative studies using the example of a multinational study about the burden of treatment

Viet-Thi Tran^{1,2,3}, Raphael Porcher^{2,3,4}, Bruno Falissard^{5,6}, Philippe Ravaud^{2,3,4,7}

¹Department of General Medicine, Paris Diderot University, Paris, France

²Centre de recherche en Epidémiologie et Statistiques, INSERM U1153, Paris, France

³Centre d'Épidémiologie Clinique, Hôpital Hôtel-Dieu, Assistance Publique-Hôpitaux de Paris, Paris, France

⁴Paris Descartes University, Paris, France

⁵Paris Sud University, Paris, France

⁶INSERM U669, Paris, France

⁷Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA

Abstract

Context

Sample size in qualitative research is ruled by data saturation, which is the point when new collected data do not enrich the study with new ideas and themes. There is no agreed method to define that point. We aimed to assess how different sampling strategies approached point of data saturation using the example of a multi-national qualitative study about the burden of treatment.

Methods

We took the opportunity of a large Web-based qualitative study about the burden of treatment involving 1,053 participants with various chronic conditions, from multiple countries, to assess the probability to identify new themes as a function of the number of patients included in the study. We used Monte Carlo simulations to determine when to stop data collection (fixed size versus use of a stopping criterion); and on which patients to include (random sampling versus selection of patients different in terms of conditions or country of residence).

Results

The median probability to identify at least one new theme for the next included subject was 32%, 24% and 12% after inclusion of 30, 50 and 100 subjects respectively. A median of 92% [IQR= 91%-93%] themes were identified with inclusion of 150 participants at random. Use of a stopping criterion was not a reliable method to reach data saturation. Purposive sampling of patients with extension of an initial sample of French patients with 5 foreign participants resulted in identification of more themes for sample size < 100 participants, than extension with French participants. However, extension of an existing sample of participants with participants having conditions less represented in the initial sample did not result in identification of more themes.

Conclusion

We assessed the impact of various sampling strategies to reach data saturation in a web-based qualitative study and obtained data saturation with a sample of 250 random participants. This may help researchers determine the right cut off between continuing the study to find new themes and stopping for futility.

1. Context

Qualitative research methods involve systematic collection, organization and interpretation of materials, usually derived from talk or observation [1]. As of today, the main criterion for adequate sample size in qualitative research is theoretical saturation, defined as the point in data collection when no new additional data that develop aspects of a conceptual category are found [2]. Operationalization of data saturation is complex as there is no agreed method to establish it [3-5]. A number of factors can affect sample size needed to obtain data saturation: 1) topic of interest [6], 2) study participants (e.g. individual factors, differences between participants) [4], and 3) methods of data collection (e.g. techniques of interviews, experience of investigators, etc.) [7, 8].

Although “in qualitative studies, sample size cannot be predicted” [9], most qualitative researchers do give a number of interviews to be conducted prior to their research to obtain funding [10] or meet ethics committees’ requests to state such numbers [11]. In those cases, sample size is determined according to studies with similar subject of interest [12], empirical recommendations or researcher’s experience [10, 13]. With increasing acceptability of the Internet and ongoing technological advances, web-based qualitative studies may represent an opportunity to gather views from patients in different settings, at low cost [14, 15]. To our knowledge, there is no guidance for sample size in such studies.

In this study, we aim to assess how different sampling strategies may approach data saturation in Web-based qualitative studies, through resampling of data generated in a previous study about the burden of treatment.

2. Methods

2.1. Materials

We used data obtained during a Web-based qualitative study exploring the components of the burden of treatment for patients with chronic conditions. English-, Spanish- and French speaking patients with at least one chronic condition, answered open-ended questions online. Participants answered: 1) a broad open-ended question intended to elicit their views on their treatments and the associated burden globally; and 2) 16 open-ended questions about different aspects of the burden of treatment identified in the literature. Participants’ answers were analyzed using “classic” content analysis with a grounded theory approach [16]: each participant’s response was read by two investigators, who independently categorized data segments into themes. For each participant, we were able to discern which themes were elicited or not, both globally and when considering only the first open-ended question.

A total of 1,053 participants answered the open ended questions from May 2013 to March 2014. Included patients came from different countries and settings, and had different conditions and treatments. In their answers, patients discussed the components, aggravating factors and consequences of the burden of treatment. More details about the study method and results are presented elsewhere [17].

2.2. Definitions

In several retrospective studies, 90% of themes were identified with the first 12 to 30 participants [4, 7] or four to five focus groups [18]. As such, authors recommended using these values for determining sample size in qualitative research. Because there is no agreed cut-off for determining data saturation as a function of the number of themes elicited, we arbitrarily defined: 1) data saturation as identifying at least 90% of all themes coded in the experimental study (independently of their nature or frequency); and 2) “frequent theme” as a

theme mentioned by more than 2.5% of patients in the original study. This limit was chosen to highlight themes mentioned by a substantial fraction of patients while keeping enough precision for thematic exploration of the concept.

2.3. Statistical analysis

We used Monte Carlo simulations to resample the experimental data to assess: 1) the probability to identify a new theme or to miss a frequent theme, as a function of the number of patients already included in the study; 2) the efficacy of two strategies to decide when to stop data collection (a priori fixed size and use of a stopping criterion); and 3) the efficacy of two strategies to decide which patients to include (purposive sampling and non-purposive sampling of participants) to elicit new themes.

All simulations were repeated 10,000 times, which ensured a precision of the estimated 95th percentile with SE=0.002. Participants were included randomly, without replacement, because they could not be considered representative of a larger population due to the purposive sampling in the experimental study. Analyses were performed using R software, version 3.1.

2.3.1. Probability to identify a new theme or to miss an important theme as a function of the number of patients already included in the study

We assessed the number of new themes elicited by each subsequent patient included in the study. From these results, we devised the probability to identify at least one new theme as a function of the number of patients already included in the study. Similarly, we estimated the probability to miss at least one “frequent” theme as a function of the number of patients already included in the study.

We studied these results in different subgroups defined by sex, level of education (primary school, high school or college) and conditions reported by the participant (diabetes, high blood pressure and rheumatologic conditions).

2.3.2. Fixed sample size versus use of a stopping criterion

We compared the efficacy of two strategies to decide when to stop data collection: 1) “a priori” fixed sample size and 2) specification of an initial analysis sample and of a stopping criterion.

To assess the efficacy to obtain data saturation of a sample of fixed size, we randomly selected samples of 15, 30, 50, 100, 150, 200 and 250 subjects among our participants. We assessed the cumulated number of themes elicited in the sample.

Second, we assessed the efficacy of a strategy described in the literature to reach data saturation [3] relying on: 1) specification of an initial analysis sample of fixed size; 2) specification of how many more interviews to conduct without new themes emerging before claiming data saturation and stopping data collection; and 3) analysis until the stopping criterion is met. We simulated this strategy by randomly selecting an initial sample of subjects among our participants and then, we iteratively added a number of random selected patients, without replacement. For each addition of participants, we assessed: 1) if the stopping criterion was met; and 2) the cumulated number of themes elicited by the sample. We tested 2 different settings: 1) initial sample of 10 + addition of iterative groups of 3; and 2) initial sample of 10 + addition of iterative groups of 10.

2.3.3. Purposive sampling versus random sampling

We compared the efficacy of two strategies to decide which patients to include in the study: 1) random sampling of participants; and 2) purposive sampling. Purposive sampling entails

active selection of the most productive sample to answer the research question based on the researcher's practical knowledge [19]. To simulate purposive sampling, we studied how data saturation would be affected by extension of an existing sample with patients different in regards of country of residence or conditions. These criteria were chosen because they were: 1) considered as the most meaningful for diversification of the sample; and 2) they represented a feasible strategy for recruitment during a web based qualitative study. First, we compared if addition of 5 foreign participants to an initial sample of French patients allowed identification of more themes than addition of other French patients. Second, we assessed if the extension of the sample by iterative addition of 5 participants, with the least represented conditions in the sample, would affect data saturation compared to addition of random participants.

2.3.4. Differences between “spontaneous answers” by participants and answers obtained after probing for specific topics

All analyses were performed using two different datasets: 1) one considering identification of themes the using the complete online survey (i.e. both the first broad open-ended question and specific questions about topics of interest identified in the literature); and 2) the other considering identification of themes using only the first open-ended question, afterwards defined as “spontaneous answers by participants”

3. Results

In our experimental study, we identified 123 themes from participants. All themes had been mentioned at least once after the 681th patient. Data saturation, that is the identification of 111 (90%) of themes, was reached after the 192th subject. Out of all 123 themes, 88 (71.5%) were elicited by more than 2.5% of participants and were considered as “frequent”; all frequent themes had been mentioned at least once after the 149th participant.

When considering only the first open-ended question, 118 themes (96%) had been elicited. Data saturation occurred only after the 976th participant. All frequent themes had been elicited at least once in the first open-ended question after the 447th participant (**Appendix 1**). There was no major influence of: sex, level of education, or conditions in the speed of discovery of new themes (**Appendix 2**).

3.1. Probability to identify a new theme or to miss an important theme as a function of the number of patients already included in the study

The median probabilities to identify at least one new theme for the next included subject, as a function of the number of participants decreased quickly and were 32%, 24% and 12% after inclusion of 30, 50 and 100 subjects, respectively. Comparatively, the median probabilities to miss at least one frequent theme were 17%, 6% and 1% after the 30, 50 and the 100 subjects, respectively (**Figure 1a and 1b**). When considering only the first open-ended question, these probabilities increased to 50%, 29% and 16% (**Figure 1a and 1b**).

3.1.1. Fixed sample size or use of a stopping criterion

We reached data saturation after inclusion of 150 participants [IQR=91-93%]. Using all data, we identified a median of 99 % [IQR= 99-100%] of frequent themes after the 100th participant. When considering only the first open-ended question, we were not able to reach data saturation with 250 participants [IQR =83-86%]. A median of 92% [IQR=89-94%] frequent themes had been spontaneously elicited at least once, after the 150th participant (**Figure 2**).

The strategy with specification of an initial sample and stopping criterion did not allow reaching data saturation for the 2 settings we predefined. Specification of an initial sample of 10 patients and iterative addition of samples of 3 participants resulted in inclusion of a median of 37 [IQR =28-43] patients eliciting a median of 72% [IQR=66%-77%] themes until stopping criterion was met. When considering a stopping criterion of 10 participants without any new theme elicited, we obtained a median sample size of 100 [IQR=90-120] patients eliciting a median of 89% [IQR=85%-91%] themes until stopping criterion was met (**Table 2 and figure 2**).

3.2. Purposive sampling versus random sampling

Purposive sampling of patients with extension of an initial sample of French patients with 5 foreign patients allowed identification of more themes than addition of French participants for initial sample size <100 patients (**Figure 3a**). Purposive sampling with iterative addition of 5 participants with conditions less represented in the sample did not improve identification of new themes as compared with the addition of random participants, whatever the initial sample of patients or study design (i.e. one open ended question only or one broad question + specific questions on predefined topics) (**Figure 3b**).

4. Discussion

Data saturation is an easy to understand concept that is difficult to ascertain in practice. It depends on the topic, the purpose of the research, the participants, the methods of data collection and analysis [20]. In the present study, we re-analyzed data from a single Web-based qualitative study involving 1,053 participants and showed that including 150 random participants ensured identification of all frequent themes and of 90% of all possible themes. Smaller studies are possible but expose to a risk of missing frequent themes. For example: a study with 100 participants allows identification of respectively 89% of all themes and 99% of frequent themes.

Literature about how saturation might be determined and for estimating adequate sample size in qualitative studies is scarce. Very few studies looked at statistical demonstration of saturation within a dataset [3, 4] and to our knowledge this is the first study to use resampling techniques to achieve this.

In this study we assessed the efficacy of two strategies to decide when to stop data collection: a priori fixed sample size and use of a stopping criterion. Use of a stopping criterion didn't reach data saturation with the two different settings we tested and there was an important variability in sample size and number of themes elicited. In the present study, it did not appear an optimal strategy for determining sample size in web based qualitative studies.

We tried to determine the benefit of purposive sampling compared to random sampling for identification of new themes. We found a slight efficacy of purposive sampling when including foreign participants; but no effect of inclusion of patients with different conditions. This may reflect that differences of burden of treatment may relate more to contextual issues (e.g. healthcare systems, cultural views of illnesses and treatments) than conditions of interest. This result underlines that purposive sampling is more complex than simple demographic stratification and should only be used for key attributes when planning initial data collection.

Addition of specific open-ended questions prompting topics of interest enhanced the ability to reach data saturation. In our experimental study, these questions were determined by a literature search and expert opinions; however, it would be possible to mimic the classic

inductive approach and determine these questions according to answers from first participants.

This study has several limitations. First, it is based on a single experimental dataset. It is likely that results would be different in other contexts, with different participants, methods of data collection and analyses. However, our study provides a proof of concept approach to data saturation for Web based qualitative studies and may help researchers plan this type of study. Second, we relied on an arbitrary definition for data saturation (90% of themes identified in the experimental study covered) or frequent themes. Because this was the first study to re-analyze by simulation the concept of data saturation in qualitative studies, we had to define limits for analysis. Finally, simulation of purposive sampling was artificial and we only considered addition of participants different than those already included in the study. “True” purposive sampling would rather imply the inclusion of patients as a function of themes and codes already analyzed [21]. This was however impossible to implement in this simulation study as we used “final data” from a previous study.

Data saturation is an elastic notion. The number of participants depends on both the purpose of research and the desired analytic level. As for all qualitative studies, new data will always add something new, but there are diminishing returns. Similarly to quantitative research where increase in sample size will always benefit statistical considerations [22], in qualitative research, it will result in identification of more fine grain themes. However, this could be limited by economic or logistical constraints: how many more participants researchers are willing to include for a single new theme? In this study, we offer the first estimate –although imperfect- of adequate sample size in web based qualitative studies. There is a need for researchers to conduct similar research with different datasets and refine the estimates in other domains of inquiry or with different participants.

5. Conclusion

Resampling of participants in a web-based qualitative study showed that data saturation was obtained with 150 random participants. This may help researchers determine the right cut off between continuing the study to find new themes and stopping for futility.

6. References

1. Malterud, K., *Qualitative research: standards, challenges, and guidelines*. Lancet, 2001. **358**(9280): p. 483-8.
2. Glaser, B. and A. Strauss, *The discovery of grounded theory: Strategies for qualitative research*. 1967, Chicago, IL: Aldine Publishing Company.
3. Francis, J.J., et al., *What is an adequate sample size? Operationalising data saturation for theory-based interview studies*. Psychol Health, 2010. **25**(10): p. 1229-45.
4. Guest, G., A. Bunce, and L. Johnson, *How Many Interviews Are Enough? An Experiment with Data Saturation and Variability*. Field Methods, 2006. **18**(1): p. 59-82.
5. Trotter, R.T., 2nd, *Qualitative research sample design and sample size: resolving and unresolved issues and inferential imperatives*. Prev Med, 2012. **55**(5): p. 398-400.
6. Charmaz, K., *Constructing grounded theory: a practical guide through qualitative analysis*. 2006, Thousand Oaks, CA: SAGE.
7. Rat, A.C., et al., *Content of quality-of-life instruments is affected by item-generation methods*. Int J Qual Health Care, 2007. **19**(6): p. 390-8.

8. Morse, J., *Styles of collaboration in qualitative enquiry*. Qualitative Health Research, 2008. **18**(1): p. 3-4.
9. Morse, J., *The adjudication of qualitative proposals*. Qualitative Health Research, 2003. **16**(6): p. 739-742.
10. Cheek, J., *The Politics and Practices of Funding Qualitative Inquiry*, in *The SAGE Handbook of Qualitative Research*, N. Denzin and Y. Lincoln, Editors. 2011, SAGE: Thousand Oaks, CA.
11. Mason, M., *Sample size and Saturation in PhD Studies Using Qualitative Interviews*. Forum: Qualitative Social Research, 2010. **11**(3).
12. Creswell, J., *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. 2002, Upper Saddle River, NJ: Pearson Education.
13. Sandelowski, M., *Sample size in qualitative research*. Res Nurs Health, 1995. **18**(2): p. 179-83.
14. Herbert, D.L., et al., *The impact of funding deadlines on personal workloads, stress and family relationships: a qualitative study of Australian researchers*. BMJ Open, 2014. **4**(3): p. e004462.
15. Ekman, A. and J.E. Litton, *New times, new needs; e-epidemiology*. Eur J Epidemiol, 2007. **22**(5): p. 285-92.
16. Lingard, L., M. Albert, and W. Levinson, *Grounded theory, mixed methods, and action research*. BMJ, 2008. **337**: p. a567.
17. Tran, V.T., et al., *Taxonomy of the burden of treatment: a multicountry Web-based qualitative study of patients with chronic conditions*. In preparation, 2014.
18. Kerr, C., A. Nixon, and D. Wild, *Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research*. Expert Rev Pharmacoecon Outcomes Res, 2010. **10**(3): p. 269-81.
19. Marshall, M., *Sampling for qualitative research*. Family Practice, 1996. **13**(6): p. 522-525.
20. Baker, S. and R. Edwards, *How many qualitative interviews are enough? Expert voices and early career reflections on sampling and cases in qualitative research*, NCRM National Centre for Research methods, Editor. 2012.
21. Coyne, I.T., *Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries?* J Adv Nurs, 1997. **26**(3): p. 623-30.
22. Altman, D.G., *Why we need confidence intervals*. World J Surg, 2005. **29**(5): p. 554-6.

Table 1: Proportion of themes elicited using random samples of participants; sample size is fixed a priori.

Study Design	Number of patients included Median - (IQR)	Proportion of themes elicited Median (%)–(IQR)	Proportion of frequent themes elicited Median (%)–(IQR)
One broad open ended question + questions on specific topics	15	54 (51-58)	70 (66-74)
	30	69 (66-72)	86 (83-89)
	50	79 (76-80)	94 (92-95)
	100	89 (86-90)	99 (99-100)
	150	92 (91-93)	100
	200	94 (93-95)	100
	250	96 (95-97)	100
Only one broad open ended question	15	30 (27-33)	40 (35-44)
	30	44 (41-47)	57 (53-61)
	50	55 (52-58)	70 (67-74)
	100	70 (67-72)	85 (83-89)
	150	77 (75-79)	92 (89-94)
	200	81 (79-84)	95 (93-96)
	250	84 (83-86)	95 (93-96)

Table 2: Proportion of themes elicited using random samples of participants; sample size is determined with an initial sample + stopping criterion.

Sampling strategy	Study design	Sample size Median–(IQR)	Proportion of themes elicited Median (%)–(IQR)
Initial sample (n=10) + Stopping if 3 consecutive participants don't elicit a new theme	One broad open ended question + questions on specific topics	37 (28-43)	72 (66-77)
	Only one broad open ended question	37 (28-46)	48 (41-54)
Initial sample (n=10) + Stopping if 10 consecutive participants don't elicit a new theme	One broad open ended question + questions on specific topics	100 (90-120)	89 (85 -91)
	Only one broad open ended question	130 (100-150)	74 (70-78)

Figure 1a: Median probability (IQR) to identify at least one new theme as a function of the number of participants already included in the study. Dark line represents a study design with both one open-ended question and specific questions on predetermined topics. Gray line represents a study design with only one open-ended question.

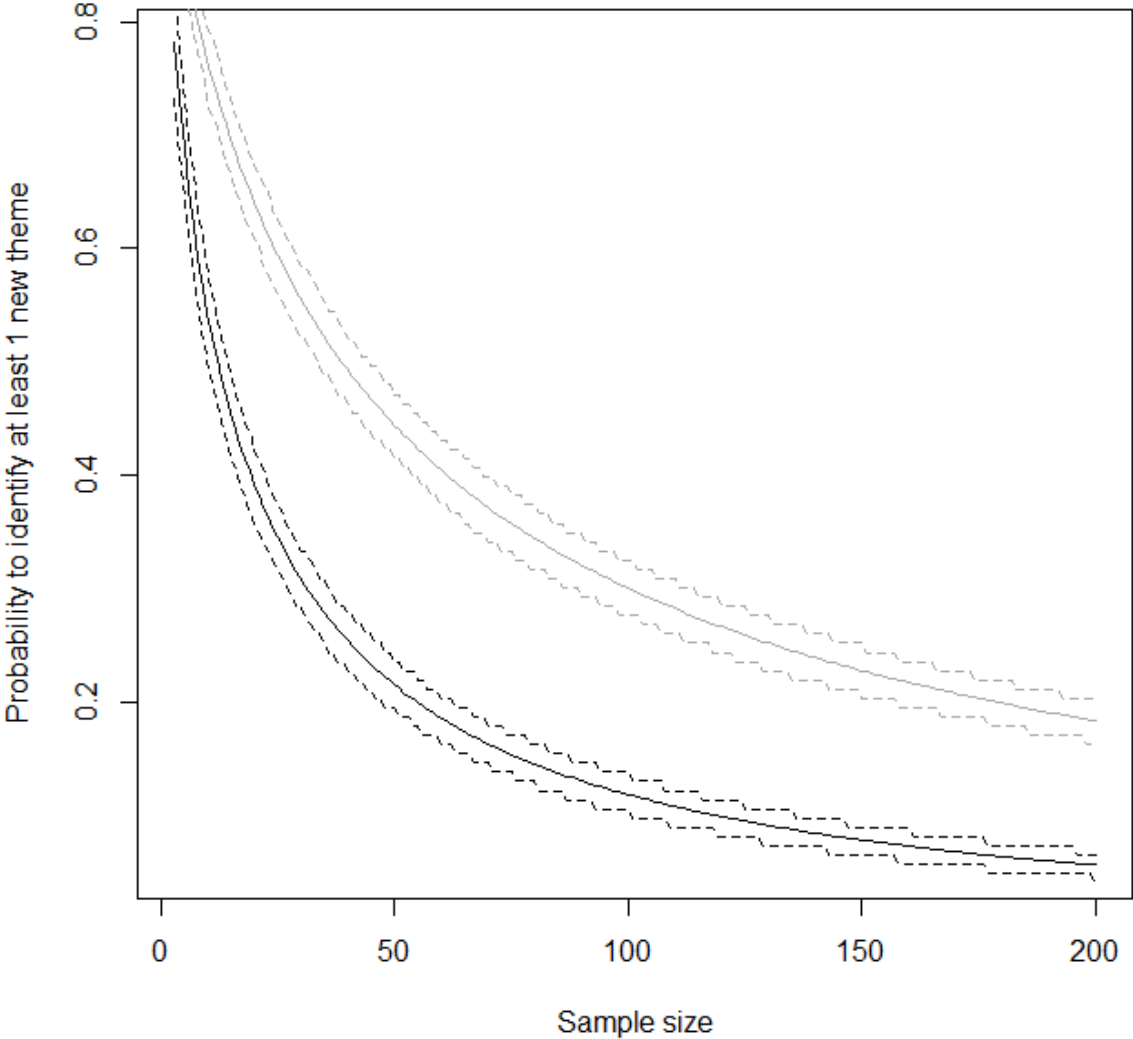


Figure 1b: Median probability (IQR) to miss at least one frequent theme as a function of the number of participants already included in the study. Dark line represents a study design with both one open-ended question and specific questions on predetermined topics. Gray line represents a study design with only one open-ended question.

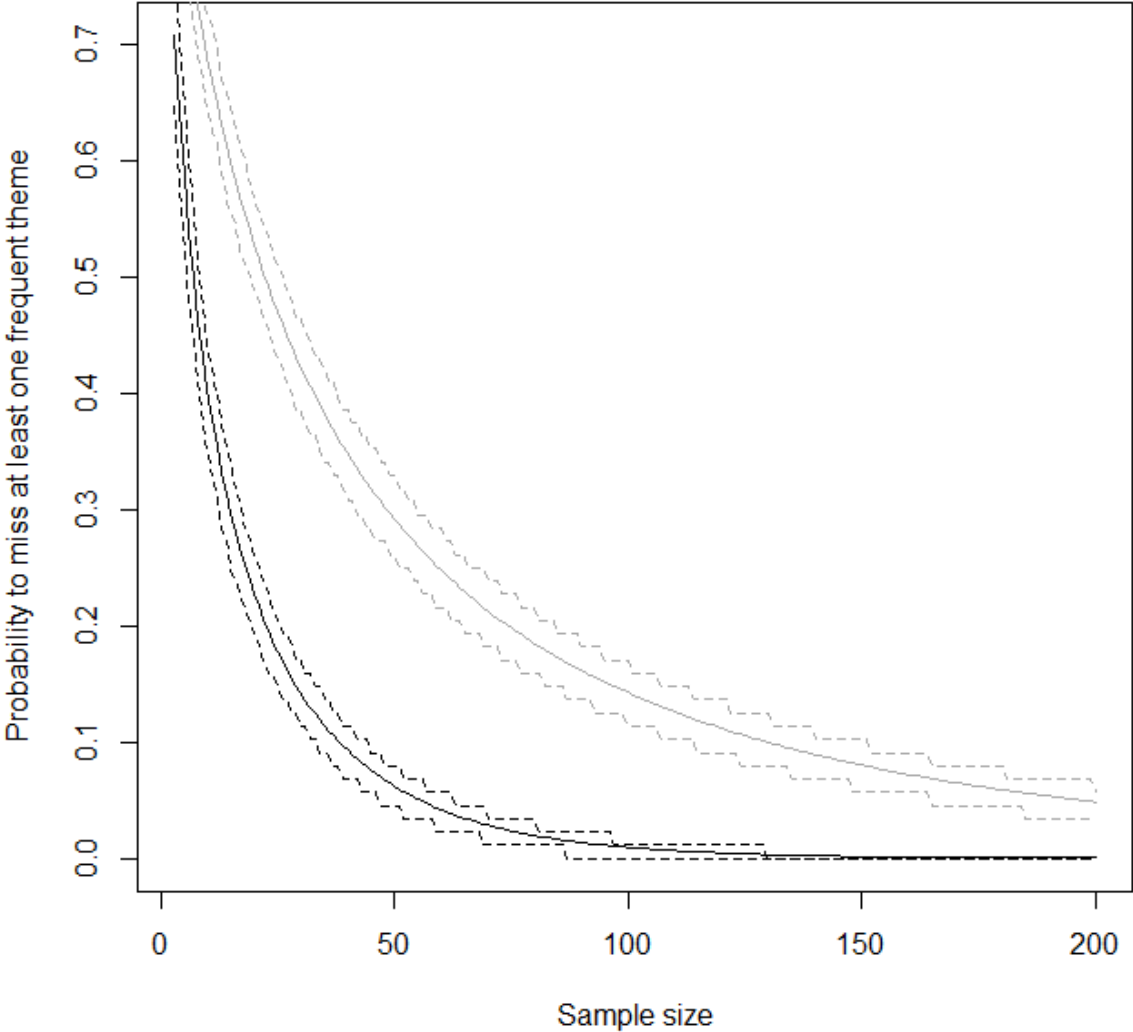


Figure 2: Proportion of themes identified with a sample determined by: 1) a fixed number of patients (black and gray box plots); or 2) with an initial sample + stopping criterion (red and green box plots). Horizontal line represents data saturation. Dark box plots represent a study design with both one open-ended question and specific questions on predetermined topics with an a priori fixed sample size. Gray box plots represent a study design with only one open-ended question an a priori fixed sample size. Red box plot represent a sampling strategy with an initial sample of 10 + stopping criterion of 3. Green box plot represent a sampling strategy with an initial sample of 10 + stopping criterion of 10.

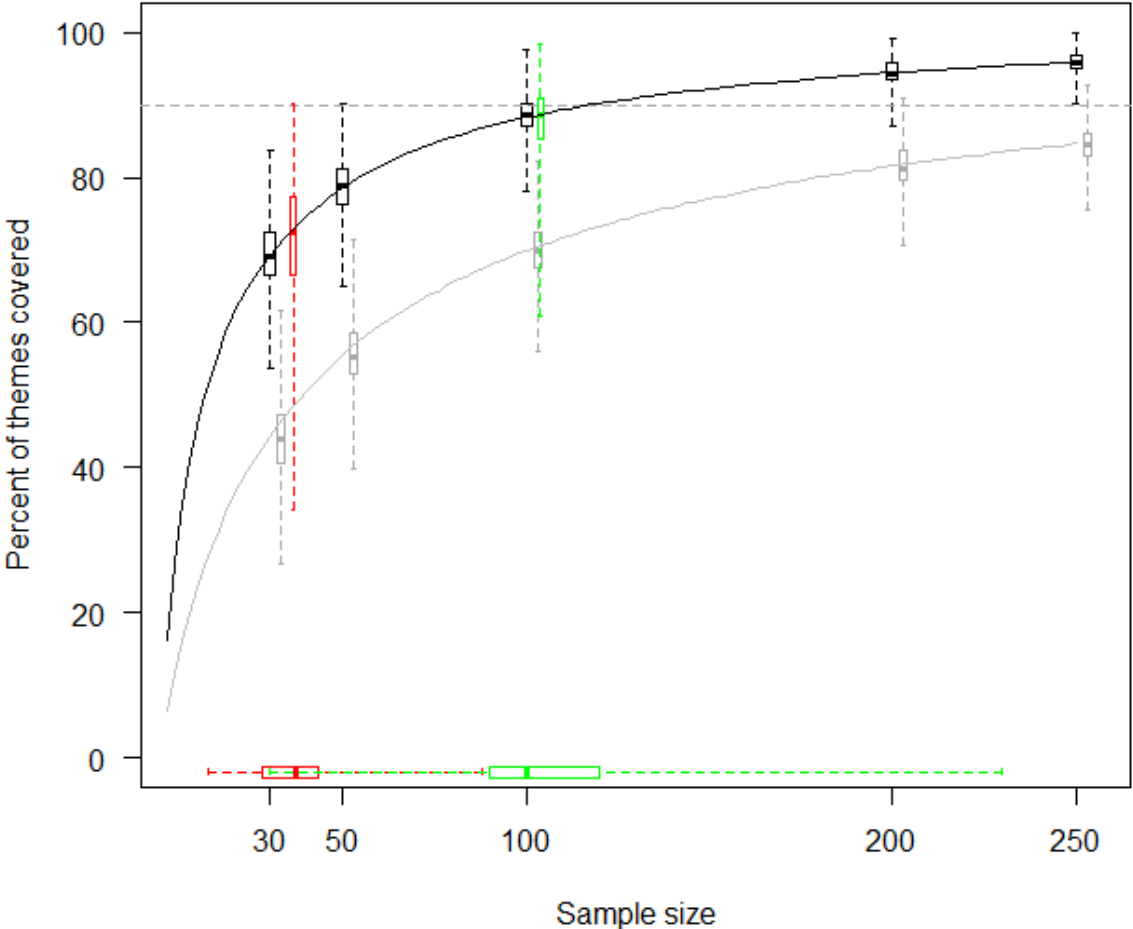


Figure 3a: Proportion of themes identified with purposive sampling for country of residence (dotted lines) versus non purposive sampling (plain lines) Purposive sampling is simulated by addition of 5 random foreign patients to an initial sample of French patients and non purposive sampling is simulated by addition of French patients. Dark lines represent a study design with both one open-ended question and specific questions on predetermined topics. Gray lines represent a study design with only one open-ended question.

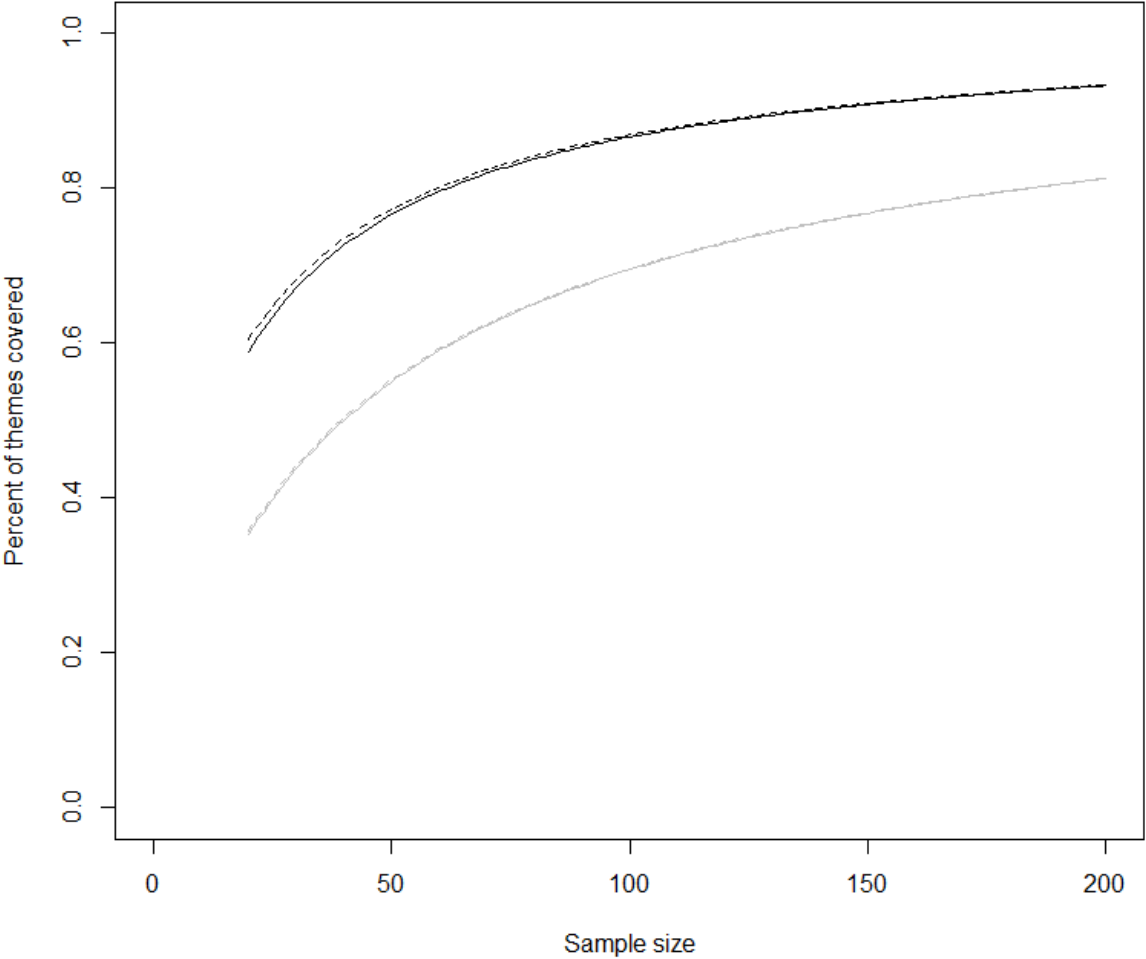
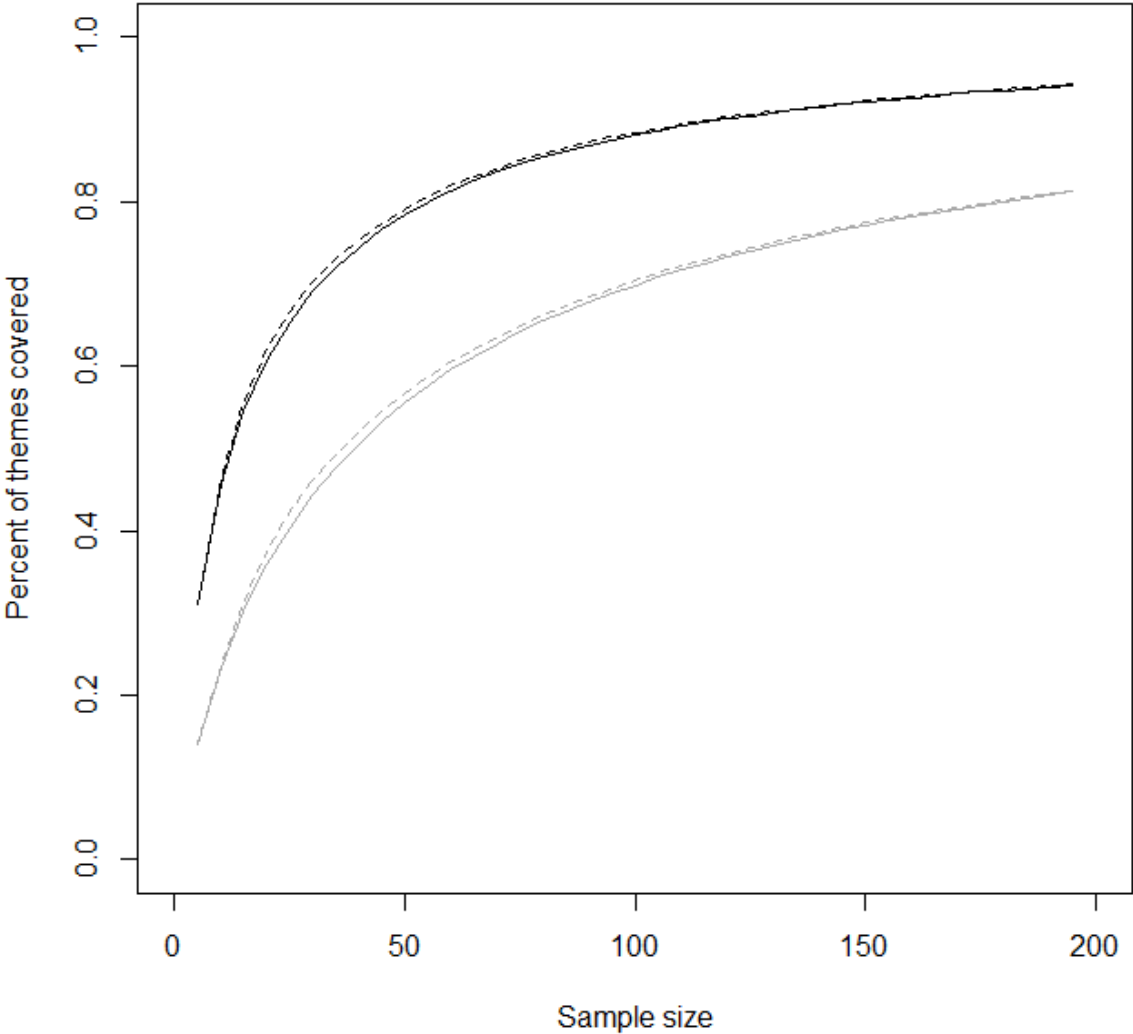


Figure 3b: Proportion of themes identified with purposive sampling for conditions (dotted lines) versus non purposive sampling (plain lines). Purposive sampling is simulated by extension of the sample by iterative addition of 5 participants with the less represented conditions in the sample. Dark lines represent a study design with both one open-ended question and specific questions on predetermined topics. Gray lines represent a study design with only one open-ended question.



7. Implications

7.1. Principaux résultats

Dans cette thèse, nous nous sommes intéressés à un concept nouveau : le fardeau du traitement. Notre travail nous a permis de développer une taxonomie du fardeau du traitement chez des patients souffrant de plusieurs maladies chroniques et de développer les premiers instruments permettant de le mesurer. Ainsi nous avons pu quantifier les relations entre fardeau du traitement, qualité de vie et observance thérapeutique.

Nous reviendrons tour à tour sur les implications de nos travaux : 1) sur le plan méthodologique, en discutant des problèmes de validité de contenu des mesures destinées à des essais internationaux et sur la conduite d'études en ligne; et 2) sur le plan thématique, en réfléchissant à l'intégration du fardeau du traitement dans la prise en charge des patients souffrant de maladies chroniques.

7.2. Implications méthodologiques : repenser le développement d'instruments de mesure provenant du patient

7.2.1. Améliorer les méthodes actuelles d'adaptation transculturelle

Sur le plan méthodologique, cette thèse nous a permis de travailler sur le développement de mesures provenant du patient. Lors de nos travaux, nous avons été confrontés aux problèmes de validité de contenu lors d'adaptations transculturelles. Le fardeau du traitement est un concept dépendant de l'environnement culturel (i.e. perception de la maladie et du traitement), familial, social et du système de soins dans lequel le patient existe. Par exemple, lors de notre premier travail, réalisé en

interrogeant des patients en France, l'aspect financier du fardeau traitement n'avait pas été mentionné lors des interviews qualitatives. En effet, notre système de santé prévoit des soins gratuits pour les patients souffrant de pathologies chroniques. En revanche, dans d'autres contextes, l'aspect financier pouvait être très contraignant pour les patients ; par exemple, 47% des patients vivant aux Etats-Unis rapportent avoir des problèmes pour payer des soins médicaux [88].

Si l'on considère que l'objectif ultime d'un instrument de mesure provenant du patient est d'être utilisé comme critère de jugement dans des essais cliniques, et que sur 37 520 essais cliniques enregistrés sur clinicaltrials.gov en 2010, 12 732 (34%) étaient internationaux [107], il est nécessaire de disposer d'instruments utilisables dans plusieurs contextes simultanément.

En utilisant les méthodes actuelles, il faudrait donc réaliser des études d'adaptation culturelle dans chaque langue cible ; ceci pouvant représenter des dizaines d'études différentes. Par ailleurs, même si les adaptations sont bien réalisées, les instruments obtenus seraient avant tout des versions « traduites et modifiées » d'un instrument initial, développé dans une population différente.

Certains auteurs ont proposé, pour résoudre ce problème, de développer simultanément l'instrument dans plusieurs contextes, en générant les items directement de manière transculturelle [75]. Quelques instruments ont été développés de cette manière, mais ces études restent minoritaires en raison de problèmes logistiques et économiques. En effet, il faudrait disposer d'investigateurs dans chaque pays [108, 109]. Par exemple, le PROQOL-HIV est un outil conçu pour mesurer la qualité de vie chez des patients infectés par le VIH. L'élaboration des items de

l'instrument a été réalisée à l'aide d'interviews qualitatives avec 152 patients dans 9 pays différents. Le temps nécessaire à la réalisation de ces interviews était de 6 mois [108].

7.2.2. Etudes utilisant des questionnaires en ligne : le futur ?

Avec le développement d'internet, le nombre d'individus connectés à internet ne cesse d'augmenter [110]. En Europe et en Amérique du Nord, on estime que la pénétration d'internet (Nombre de personnes utilisant internet/nombre d'habitants dans ces régions) est > 70% (83% en France, 89% au Royaume Uni, 87% aux Etats-Unis et 90% au Canada). Dans ces pays, les études en ligne pourraient permettre de palier au problème précédemment exposé, en facilitant le recrutement de patients distant géographiquement, à faible coût. Notre thèse a permis de montrer la faisabilité de cette approche avec deux études originales par leur méthodologie en ligne.

L'utilisation d'internet a permis de recruter un grand nombre de patients en peu de temps : lors de l'adaptation du TBQ en anglais, nous avons pu recruter 800 patients en 2 mois; et lors de l'étude qualitative, nous avons obtenu les verbatim de 1053 patients en 10 mois, c'est-à-dire 4,2 fois plus rapidement que pour le PROQOL-HIV [108]. Outre le gain de temps, l'utilisation d'une plateforme en ligne a permis d'obtenir des réponses de meilleure qualité, en utilisant des systèmes automatiques pour vérifier la cohérence des données. Dans la littérature, il est également décrit que l'utilisation d'internet permet d'assurer un meilleur sentiment d'anonymat du participant, réduisant un biais de désirabilité sociale dans ses réponses [111]. Par exemple, lors de l'adaptation du TBQ en Anglais, les patients ont insisté pour intégrer un item sur les problèmes qu'ils pouvaient avoir lors de leurs relations avec les médecins. Cet item n'avait pas été

identifié au cours des entretiens qualitatifs réalisés face-à-face avec le patient, en France, peut-être parce que ces patients étaient moins enclins à discuter de cela avec les investigateurs. Un autre avantage des études en ligne est la capacité des patients à répondre de chez eux. Habituellement, les patients participant à des recherches doivent se rendre dans des centres de recherche. Les études en ligne permettent ainsi de recruter des patients vivant loin des centres de recherches habituels ou ceux dont la mobilité pourrait éventuellement être un frein à la participation. L'utilisation de la plateforme de PatientsLikeMe lors de l'adaptation du TBQ a également permis d'inclure des patients motivés par la recherche. Lors de leur participation, les patients n'ont pas uniquement rempli les questionnaires que nous leur proposons mais ont également cherché à améliorer l'instrument, en proposant par exemple des reformulations des items qu'ils avaient eu du mal à comprendre [112].

Ces études ne sont pas dénuées d'inconvénients. Dans nos travaux, nous avons recruté des patients plus jeunes avec un niveau d'éducation plus élevé. Ces résultats concordent avec ceux d'autres études en ligne [85, 86, 113]. Bien que cela introduise indubitablement un biais dans la recherche, celui-ci est simplement différent de celui induit lors de l'utilisation de questionnaires papiers [114, 115]. La vraie question est: à quel point ces deux estimations sont fiables, c'est à dire éloignées de l'état réel du phénomène à mesurer ?

Avec le développement et la démocratisation de l'accès aux nouvelles technologies, il est probable que, dans le futur, la plupart des individus seront familiers avec l'utilisation de ces technologies [95]. A ce moment, il sera alors même envisageable

que les études en ligne soient moins biaisées que les études avec un recrutement « classique » face-à-face, à l'aide de questionnaires papiers.

7.3. Implications pour les patients : repenser la prise en charge des patients avec des maladies chroniques

Nos travaux nous ont permis de mieux comprendre le concept du fardeau du traitement. Lors de notre étude qualitative, nous avons mis en évidence les différentes tâches imposées aux patients et les facteurs aggravants, personnels, structurels ou situationnels aggravant le fardeau du traitement associé à ces tâches. Dans de nombreuses situations, le fardeau du traitement est inévitable : de multiples traitements médicamenteux sont souvent nécessaires pour contrôler les symptômes ou prévenir des exacerbations. Cependant, dans d'autres situations, il peut être réduit.

La réduction du fardeau du traitement comprend trois niveaux : 1) un niveau « global » portant sur l'amélioration des recommandations de pratiques cliniques ; 2) un niveau « régional », portant sur l'organisation des soins et 3) un niveau « local » portant sur les changements de mentalité des médecins et patients.

7.3.1. Développer des recommandations adaptées à la multimorbidité

Actuellement, la plupart des recommandations de pratique clinique sont développées pour des maladies données et ne permettent pas de prendre en charge des patients complexes [116]. A l'heure actuelle, les recommandations sont fondées sur des preuves ne correspondant pas aux patients à qui elles sont appliquées [27] et ne prennent pas en compte la multimorbidité [31]. Il est donc indispensable de repenser les recommandations afin qu'elles soient plus proches de la réalité. Même s'il est impossible de produire des recommandations couvrant toutes les situations

rencontrées en clinique, les futures recommandations devraient intégrer les conduites à tenir pour les « clusters de maladies chroniques » les plus fréquents [15, 117, 118]. Ceci pourrait permettre de simplifier la prise en charge des patients « types » multimorbides.

7.3.2. Lutter contre les prescriptions inutiles

Par ailleurs, il est nécessaire de lutter contre la multiplication d'interventions ou d'examens peu efficaces, prescrits pour « rassurer » les médecins (en éliminant des causes improbables de maladies, ou pour se couvrir), voire pour leur assurer un revenu [31]. La prescription de tests inutiles augmente le risque de faux positif, d'examens ou traitements supplémentaires [119], contribuant au fardeau du traitement.

Il est donc essentiel d'identifier et de recenser ces interventions inutiles, afin d'établir des recommandations de traitements ou examens à « ne pas prescrire ». Aux Etats-Unis et au Canada, cette initiative s'intitule « Choosing Wisely » où chaque société savante nationale publie une liste d'interventions fréquemment prescrites, offrant peu de bénéfices aux patients [22].

Sur le plan pharmacologique, chez les patients âgés de plus de 65 ans, de nombreuses initiatives existent déjà, tels les critères de Beers [120] ou STOPP [121, 122], fondés sur les traitements dont la balance bénéfice risque décroît chez les patients âgés. L'utilisation de ce type de critères a montré une réduction des prescriptions et de la iatrogénie chez les patients âgés [123]. Une piste afin d'améliorer la prise en charge des maladies chroniques en pratique pourrait être le développement de critères similaires adaptés à la multimorbidité.

7.3.3. Dépasser les recommandations pour individualiser la médecine

Les résultats de la recherche –sous la forme des recommandations- sont nécessaires mais non suffisants pour bien prendre en charge les malades. Les traitements recommandés ne représentent pas toujours une option pour un patient donné. En effet, les patients sont inégaux face aux demandes imposées par leurs soins : deux patients avec les mêmes pathologies et traitements pouvaient vivre de manière différente, avoir des attentes différentes vis-à-vis de leurs traitements. Il faut donc dépasser l'application « à la lettre » des recommandations de pratique cliniques « One size fits all » pour personnaliser la prise en charge des patients à leurs contextes et capacités, c'est-à-dire pratiquer une « Minimally Disruptive Medicine ». A titre d'exemple, les recommandations américaines préconisent un traitement par statine chez des patients diabétiques entre 40 et 75 ans, quel que soit leur risque cardiovasculaire. Cependant, l'information éclairée du patient sur le rapport entre bénéfices, risques, et fardeau du traitement était associé à une réduction de 70% de leur désir à suivre ce traitement [124]. Ainsi, suivre cette recommandation sans prendre en compte les vues du patient pourrait être à l'origine d'une inobservance de ce traitement. Le médecin face à une situation complexe doit donc savoir s'éloigner des recommandations et suivre ce que son sens clinique lui dicte de faire [25]. Pour arriver à cet objectif, il faudra à la fois : 1) un changement de politique de santé, afin que le paiement à la performance puisse prendre en compte ces compromis entre recommandations et capacités du patient [25] ; et 2) un changement de mentalité des médecins pour sortir du sentiment de culpabilité lorsqu'ils ne suivent pas les recommandations « à la lettre » [116].

Prendre en compte les vues des patients est souvent plus complexe qu'il n'y paraît. Pour que cela ait du sens, il faut que les patients aient une vision fiable des données de la science pour interpréter les bénéfices et risques associés à chaque intervention. En effet, les patients, comme les médecins, peuvent souvent avoir une vision biaisée de ce qui est important pour eux. Par exemple, un diabétique sur 4 considère que l'hémoglobine glyquée HbA_{1c} est aussi importante que la survenue de complications majeures de leur maladie [125], fait aujourd'hui infirmé par la recherche. En clinique, l'utilisation d'aides à la décision pourrait aider les patients et cliniciens à prendre leurs décisions avec toutes les données en main.

Pour mieux suivre les patients avec des maladies chroniques, il faut donc valoriser une médecine adaptée tenant compte à la fois des données de la science, des préférences du patient et de ses capacités, c'est-à-dire pratiquer une « Evidence Based Medicine ».

7.3.4. Changer le modèle de consultation

La multimorbidité doit également amener à repenser le modèle de consultation. Actuellement, les consultations suivent encore le modèle proposé par Stott et Davis en 1970 [126] et sont centrées sur des problèmes aigus. Cette consultation propose 4 temps : 1) prise en charge du problème pour lequel le patient consulte ; 2) prise en charge de problèmes de fond ; 3) planification du suivi futur prenant en compte à la fois des « objectifs de sécurité » et les attentes du patient; et 4) promotion de la santé « opportuniste ». Pour des patients multimorbides, cette segmentation arrive à ses limites car le temps de consultation dédié au suivi des maladies, à l'évaluation de l'observance et aux objectifs de prévention secondaire, peut rapidement augmenter et prendre le pas sur d'autres problèmes du patient. Par exemple, dans nos travaux, les

patients étaient souvent frustrés que leurs médecins passent plus de temps à « cocher des cases » qu'à les écouter.

Pour palier à ces limites, Muth et al. proposent de changer les temps de consultations habituels chez les patients multimorbides. Selon les auteurs, la consultation devrait prendre en compte : 1) l'évaluation des interactions entre traitements et maladies ; 2) l'évaluation des priorités et des préférences du patient et 3) l'individualisation de la prise en charge et du suivi, en particulier en prenant en compte le fardeau du traitement des patients (**Table 5**). Pour cela, ils proposent un « fil d'Ariane pour guider le médecin dans le labyrinthe de la prise en charge des maladies chroniques » [127].

Assess potential interactions – the patient's conditions and treatments, constitution and context

- Keep a list of all current conditions, assess their severity and impact, and review the medication currently taken.
- Actively monitor for signs of anxiety, distress and depression, or cognitive dysfunction, including problems of addiction and non-specific signs or symptoms such as sleeping problems, loss of appetite, or hydration problems.
- Elicit and consider social circumstances, financial constraints, living conditions and social support, health literacy, functional autonomy, and coping strategies.
- List other physicians and therapists involved in the patient's care and assess overall treatment burden.

Elicit preferences and priorities – the patient's most and least desired outcomes

- Elicit preferences for generic health outcomes, such as survival, independence, pain, and symptom relief including palliative care needs, and be aware of your own (implicit) preferences, as they may not be the same as the patient's.
- If applicable, consider preferences of informal caregivers or family.
- Agree on a realistic treatment goal with the patient (and patient's caregiver if appropriate).

Individualize management to reach the negotiated treatment goals

- Weigh up whether the expected benefits of treatment (and prevention) outweigh the likely downsides and harms, given the individual patient's risk level and preferences.
- Assess the incremental and combined treatment burden of the patient (and caregiver, if applicable).
- Consider self-management according to the patient's needs and capabilities.
- Provide instructions for safety-netting such as symptoms of side effects and recommendations about the appropriate management.
- Agree with the patient on the schedule for follow-up visits to evaluate goal attainment and re-assess interactions.
- Consult other health care providers and informal caregivers who are involved with the

patient. Ideally, all health care providers involved are informed about treatment decisions or have access to information.
--

Table 5: Principes d'Ariane servant à guider le médecin dans le labyrinthe de la prise en charge des patients multimorbides. Adapté de la source [127].

Bien que pertinents, ces principes ne pourront être mis en pratique qu'en donnant plus de temps aux médecins. Actuellement, la durée de consultation varie entre 17 et 21 minutes [128]. Ce temps, inextensible, est déjà souvent largement occupé par le problème pour lequel vient consulter le patient, laissant peu de temps au suivi des maladies « de fond ». Pour libérer du temps dans les consultations des médecins sans en sacrifier la qualité, une piste serait de repenser le rôle du médecin et des paramédicaux dans la prise en charge du malade chronique. Par exemple, dans les « Patient Care Team », certains aspects du suivi habituellement réalisés par le médecin sont délégués à des soignants paramédicaux [129] dans un travail d'équipe. Ces soignants peuvent alors assumer 3 fonctions différentes : 1) un « supplément », en déchargeant le médecin de tâches qu'il pourrait réaliser, mais de manière peu efficace (i.e. injections, etc.) ; 2) un « complément », en réalisant des tâches que les médecins ne réalisent pas, faute de temps ou de moyens (i.e. conseil, consultations motivationnelles, etc.) ; et 3) un « substitut », lorsque ces soignants s'occupent de tâches habituellement réalisées par des médecins (i.e. diagnostic ou traitement).

7.3.5. Adapter la consultation à la complexité

Cependant, il ne faut pas que ces nouvelles organisations augmentent le fardeau du traitement des patients. Tous les patients multimorbides ne sont pas complexes! Il est donc important d'adapter la prise en charge à la complexité du patient [116] : si les cliniciens devraient pouvoir orienter les patients complexes vers des consultations

spécifiques [25], éventuellement dédiées à la résolution de problèmes spécifiques [130], ils devraient également pouvoir réduire le suivi d'autres patients, plus simples, sans problèmes, ayant bien intégré leurs maladies et traitements.

Pour cela, il est nécessaire de disposer d'indicateurs fiables de la complexité des patients afin de leur offrir la prise en charge la mieux adaptée. La plupart des indicateurs proposés dans la littérature reposent sur un décompte de maladies [11] ou de traitements [131]. Cependant, toutes les maladies chroniques ne se valent pas. Pour mieux définir la complexité d'un patient, une piste proposée par Christensen et al. est de repenser la classification des maladies en fonction de leur prise en charge [132] .

7.3.6. Adapter la typologie des maladies chroniques à leur prise en charge

Pour définir de nouvelles organisations, en particulier pour identifier les patients pouvant bénéficier de prise en charge simplifiée ou à l'inverse les patients ayant besoin d'une attention particulière, Christensen et al. propose de reclasser les maladies en :

- « Rule based diseases » où maladies dont la prise en charge est codifiée (par exemple: diabète de type I ou II, hépatite chronique, hypertension artérielle, etc.). La prise en charge optimale de ces maladies peut être réalisée par un praticien seul aidé éventuellement de personnels.
- « Intuitive diseases » où maladies dont le diagnostic ou la prise en charge est mal codifiée, fondée sur de l'expérience clinique (par exemple : douleur chronique, dépression, obésité, etc.). Elles requièrent l'intuition de personnels de haut niveau performants et nécessitent des tâtonnements pour arriver à une prise en charge efficace. La prise en charge optimale de ces maladies nécessite un travail d'équipe entre plusieurs médecins.

Cette nouvelle typologie, adaptée aux organisations précédemment présentées, permettrait de mieux orienter les patients dans le système de soins, offrant à chaque patient une organisation adaptée à ses problèmes.

7.3.7. Améliorer la coordination entre soignants

Lorsque les patients sont complexes, leur prise en charge nécessite souvent une fragmentation entre plusieurs intervenants. Le modèle actuel consiste à adresser le patient à un autre médecin spécialisé dans le problème donné ; médecin qui va alors prescrire des interventions ou initier un suivi avec ce patient. Ce modèle pose problème s'il existe une perte d'information entre soignants ou si les vues des médecins ne sont pas coordonnées et donc à l'origine d'un conflit dans les priorités et traitements proposés aux patients. Safran et al. propose la métaphore d'un corps de ballet : pour créer une belle chorégraphie synchronisée, chaque danseur doit être à la fois attentif à ses propres mouvements et à ceux des autres danseurs, ajustant continuellement sa position. Selon l'auteur, en médecine, les choses devraient être semblables ; chaque médecin devrait ajuster ses décisions en fonction des autres intervenants. Malheureusement, la formation des médecins insiste davantage sur leurs performances personnelles que sur leur capacité à travailler en équipe [133]. Au final, les médecins n'ont pas de culture de travail d'équipe et cherchent rarement à coordonner leurs efforts autour d'un patient donné. Et, quand bien même ils voudraient le faire, ils se retrouvent rapidement limités par la faible performance des systèmes d'information [129].

Sans modifier ces deux points, l'implémentation de systèmes « intégrés », où les patients sont pris en charge par des équipes pluridisciplinaires, sera compromise et les

soins resteront fragmentés. Ceci doit se faire parallèlement au développement d'un lien entre ville et hôpital. Ce lien doit aller de la ville vers l'hôpital avec la prise en compte lors d'une hospitalisation du contexte du patient [134]; mais également de l'hôpital vers la ville avec le développement de stratégies de « translational care » destinées à éviter les retours en hospitalisation [135].

7.3.8. S'adapter au patient du futur

Pour améliorer la prise en charge des patients, il faudra également que les médecins s'adaptent aux patients d'aujourd'hui –et de demain-. Les patients vivent dans un monde connecté où l'information est facilement accessible : 80% des internautes ont déjà cherché une information relative à leur santé en ligne [136]. Ceci est illustré par l'émergence de grandes communautés en ligne où les patients peuvent échanger et discuter avec d'autres patients sur leurs symptômes et prise en charge. Dans ces communautés, les patients apprennent les traitements potentiellement efficaces pour traiter leurs pathologies par d'autres patients, et peuvent décider de les prendre – parfois sans l'avis de leurs médecins- [137]. Ceci marque une rupture avec le modèle classique de soins où le médecin décidait des traitements utiles. Désormais, les patients s'informent et débutent par eux-mêmes des interventions en lesquelles ils croient (en particulier des régimes, des programmes d'exercices physiques, etc.). Le patient va alors adapter l'intervention à ses capacités : optimiser sa prise en charge en fonction de ses autres rôles et capacités, c'est-à-dire en fonction du fardeau du traitement. De son côté, le médecin va devoir assumer un nouveau rôle, passant de prescripteur à conseiller : l'intervention est-elle efficace ? Adéquate ? Bien conduite ? Le patient renonce-t-il à des soins utiles pour réaliser d'autres interventions dont il aurait entendu parler ?

La technologie peut également être utilisée par le médecin pour améliorer le suivi des patients. Par exemple, le simple fait de recevoir un SMS tous les jours améliore significativement l'observance de patients souffrant du VIH et augmente la proportion de patients avec une charge virale indétectable [138]. Le développement des smartphones a offert l'opportunité aux patients de mesurer différents paramètres physiologiques (tels que le pouls, le nombre de pas, l'activité physique, etc.) pour un coût extrêmement faible [139]. Des appareils plus perfectionnés, sans fils, légers, existent déjà pour surveiller de manière fiable et continue de nombreux autres paramètres physiques sans accroître les contraintes pour le patient [140]. Environ 1 américain sur 4 utilise déjà un outil connecté pour surveiller sa santé [141]. A l'heure actuelle, ces moyens sont encore des gadgets, mais l'évolution des techniques permettra probablement, dans le futur, de réaliser une partie des tâches que les patients considèrent comme pénibles (i.e. auto mesure de la tension artérielle, surveillance glycémique, bilans biologiques, etc.) de manière plus transparente pour eux, via des outils connectés.

La technologie pourrait également être utile pour faciliter la communication entre patients et médecins. Même si une méta-analyse Cochrane évaluant l'impact de l'utilisation d'e-mails pour faciliter la communication entre médecins et patients n'a pas montré de différence en termes d'utilisation des ressources de soins [142], il est possible que ce mode de communication rapide et direct avec le médecin ait un impact positif sur le fardeau du traitement. Des recherches en ce sens sont nécessaires. En ce qui concerne la communication entre médecins, l'utilisation des nouvelles technologies pourrait aussi représenter une alternative fiable à l'adressage à un autre médecin. Par exemple, la concordance entre le diagnostic d'un dermatologue à partir

de photographies prises à l'aide d'un téléphone portable et celui donné au cours d'une consultation face-à-face avec le patient est excellente ($\kappa=0.91$) [143] : l'utilisation de telles interventions pourrait donc épargner aux patients des consultations, temps de rendez-vous, temps d'attente et améliorer leur prise en charge. L'intégration de nouvelles technologies de communication permettant l'envoi rapide, facile et sécurisé de données cliniques, entre médecins pourrait donc réduire le besoin d'avis « physiques » de spécialistes, et ainsi réduire le fardeau du traitement des patients.

Le patient d'aujourd'hui et de demain vit dans un monde connecté. Avec douze milliards de smartphones à travers le monde [141], et le développement effréné d'applications mobiles dédiées à la santé, il est crucial que les médecins et chercheurs intègrent au mieux ces technologies dans la prise en charge de leurs patients, pour réduire le fardeau du traitement.

8. Perspectives

Notre thèse amène également un certain nombre de développements à plus court terme. Nous présenterons dans cette partie quelques pistes de travail.

8.1. Perspectives méthodologiques

8.1.1. Comparer les méthodes de recueil de données dans les études qualitatives

Lors de notre thèse, nous avons réalisé une étude qualitative en ligne impliquant 1 053 participants dans 34 pays différents. A notre connaissance c'est la plus grande étude qualitative de ce type. Nos travaux ont permis de montrer la faisabilité de cette approche et donner les premiers outils pour estimer la taille d'échantillon dans des études similaires. Cependant, des questions restent en suspens. Notamment, la qualité des résultats en utilisant cette méthode par rapport à une approche plus classique avec des interviews ou focus-groups « face-à-face ». Rat et al. a montré que le mode de recueil des données lors d'études qualitatives influait significativement sur les thèmes identifiés [144]. D'ailleurs, nous avons supposé que certains thèmes avaient été mieux décrits lors de l'étude en ligne où les patients avaient un sentiment d'anonymat, que lors des interviews qualitatives menées face-à-face. Une perspective de travail futur pourrait donc être la comparaison entre les thèmes identifiés lors d'études en ligne et classiques.

8.1.2. Améliorer l'évaluation de la validité de mesures provenant du patient

La réalisation de larges études qualitatives internationales, dont les échantillons se rapprochent de ceux d'études « quantitatives », est faisable à faible coût. Il serait donc

possible, au cours d'une étude mixte, de comparer les résultats provenant d'études qualitatives et les mesures « numériques » provenant d'instruments cherchant à évaluer le même concept. Cette méthode, à la limite entre recherche qualitative et quantitative, pourrait être considérée comme une extension de la « double interview » utilisée lors du développement d'instruments. La double interview consiste à demander au participant de remplir l'instrument seul pour commencer; puis, dans un second temps, d'expliquer à l'investigateur les raisons pour lesquelles il a répondu comme il a répondu. Jusqu'alors, cette méthode n'avait été réalisée qu'avec de petits effectifs (<20) pour des raisons de faisabilité [66]. Nous proposons de saisir l'opportunité de grandes études qualitatives en ligne pour généraliser cette méthode. Si nous considérons l'interview comme une manière d'évaluer un phénomène, il devrait exister une concordance entre l'évaluation du phénomène par cette méthode et par une mesure provenant du patient. Ainsi, nous pourrions évaluer s'il existe un seuil de l'instrument de mesure en dessous duquel un phénomène, même s'il est perçu par le patient, n'est pas mentionné dans une interview ; ou si à l'inverse, des éléments mentionnés dans une interview seraient évalués à « 0 » par l'instrument. Enfin, nous pourrions nous assurer qu'il n'y avait pas, dans l'interview, de points non abordés dans l'instrument. Cette étude de concordance donnerait donc une évaluation de la validité de l'instrument.

8.1.3. Estimer le nombre de sujets nécessaires dans les études qualitatives

Nous aimerions poursuivre notre travail sur la taille d'échantillon dans les études qualitatives en développant un modèle permettant d'estimer, à partir des résultats des premières interviews ou focus groups, le nombre final de thèmes et le nombre de participants à inclure. Ce projet part de l'hypothèse d'une tendance similaire des

courbes présentant le nombre de thèmes cumulés identifiés en fonction du nombre de participants [98, 105]. En ré-analysant des études qualitatives, nous pourrions, dans un premier temps, vérifier cette hypothèse. Puis, si elle s'avérait exacte, nous pourrions développer un modèle prédictif de la taille d'échantillon en fonction de données préliminaires. Ce modèle pourrait alors être mis à disposition des chercheurs afin qu'ils puissent estimer le nombre de sujets nécessaires en fonction des premiers résultats de leurs études, sous forme d'un outil en ligne. Les données utilisées permettraient de préciser le modèle. Ainsi, nous pourrions imaginer obtenir une « méthode de calcul » du nombre de sujets nécessaires dans les études qualitatives en fonction des premières interviews.

8.2. Perspectives de recherches sur le thème du fardeau du traitement

8.2.1. Evaluer la part du fardeau du traitement portée par le patient et ses aidants

Dans nos travaux, nous nous sommes attachés à décrire le fardeau du traitement du point de vue des patients. Cependant, Corbin et Strauss avaient déjà identifié que dans les familles des patients, les tâches relatives à la santé ou la vie familiale étaient réparties entre les membres du foyer à cause de la maladie [37]. Ainsi, la capacité d'un patient donné à accomplir l'ensemble des tâches nécessaires à se soigner peut être compensés par l'aide de ses proches. On estime qu'environ 90% des aidants des patients avec des maladies chroniques sont la famille (ou des proches non rémunérés) [145]. Nous aimerions explorer comment le fardeau peut être réparti entre le patient et ses aidants. A quel point y a-t-il un transfert de ce fardeau ? Les attentes des deux parties sont-elles similaires ? Pour ce faire, nous pourrions adapter notre mesure du

fardeau du traitement afin que le patient et son aidant identifient, ensemble, comment ils se sont répartis les tâches imposées par le traitement. Les résultats de cette étude nous permettrait d'identifier des interventions permettant d'optimiser le traitement des patients, sans aggraver la charge des aidants, afin d'améliorer la capacité du « groupe ».

8.2.2. Identifier ce que les patients aimeraient modifier dans leur prise en charge

Pour réduire le fardeau du traitement, les médecins doivent mieux adapter ce qu'ils proposent aux patients à leurs capacités et à leurs attentes. Dans une population de patients homogènes, de part leurs traitements et maladies, nous aimerions identifier ce que les patients aimeraient modifier dans leur prise en charge « s'ils avaient une baguette magique ». Les résultats de ce travail pourraient permettre aux médecins de mieux comprendre quelles sont les choses qui sont les plus difficiles à vivre par leurs patients. Dans un second temps, nous pourrions comparer ces « interventions », définies par les patients, avec leur faisabilité, évaluée par leurs médecins. Ce projet, pouvant être répété dans plusieurs pathologies données, pourrait ainsi permettre de faciliter la priorisation des interventions.

8.2.3. Evaluer l'impact de consultations dédiées à la multimorbidité sur la morbidité et la qualité de vie des patients chroniques.

Dans la partie précédente, nous avons discuté sur de possibles nouvelles organisations pour améliorer le suivi des patients avec des maladies chroniques. L'implémentation d'une consultation dédiée à un sujet spécifique complexe est souvent plus efficace que lorsque ce sujet est abordé lors de temps courts dans plusieurs consultations [130, 146].

C'est pourquoi nous aimerions évaluer l'impact d'une consultation standardisée par le médecin traitant, centrée autour de la multimorbidité. Cette étude serait un essai contrôlé randomisé. Les patients concernés seraient complexes (c'est-à-dire présentant >3 maladies chroniques, ou >2 maladies dont une psychiatrique, ou >10 traitements médicamenteux). L'intervention serait une consultation standardisée basée sur le fil d'Ariane proposé par Muth et al. [127] et permettrait au médecin de personnaliser et d'adapter la prise en charge de chaque patient en fonction de ses attentes et du fardeau du traitement. Le contrôle serait une prise en charge « habituelle » par le médecin traitant. Les critères de jugement de cette étude pourraient être l'observance thérapeutique, la survenue de complications ou d'hospitalisation et la qualité de vie des individus.

9. Conclusion

Une tempête parfaite survient lorsque plusieurs facteurs ou évènements –aucun d’entre eux n’étant particulièrement dévastateurs- se cumulent en une force « catastrophique » [147]. Ceci pourrait être le cas pour le suivi des patients avec des maladies chroniques, qui:

- est encore centré sur la maladie, peinant à concilier le dilemme entre suivi spécialisé pour offrir un niveau de soins techniquement adapté à chaque maladie ; et prise en compte des malades dans toutes leurs dimensions (avec leurs différentes maladies, contexte personnel, etc.).
- ne tient pas compte la multimorbidité
- ne tient pas compte de l’évolution des patients, qui ont accès à une information foisonnante de qualité variable et qui initient par eux même des programmes thérapeutiques divers (régimes, exercices physiques, voire même traitements).
- ne profite pas (ou peu) des évolutions technologiques et des outils connectés utilisés quotidiennement par les patients et médecins, sauf lors des soins.

Ces facteurs contribuent à un suivi des patients chroniques fragmenté, mal coordonné, à l’origine de surdiagnostic, de surtraitement et de fardeau du traitement.

Il est donc nécessaire de repenser la prise en charge des maladies chroniques, en plaçant le patient au centre des soins. La première étape pour arriver à une médecine adaptée au contexte des patients et ayant le moins d’impact possible sur leurs vies quotidiennes était de disposer d’outils permettant d’évaluer le fardeau du traitement.

La prochaine étape consistera à développer et évaluer de nouvelles organisations du suivi des patients multimorbides efficaces et minimisant le fardeau du traitement.

10. Bibliographie

1. May, C., V.M. Montori, and F.S. Mair, *We need minimally disruptive medicine*. BMJ, 2009. **339**: p. b2803.
2. Banting, F.G., W.R. Campbell, and A.A. Fletcher, *Further Clinical Experience with Insulin (Pancreatic Extracts) in the Treatment of Diabetes Mellitus*. Br Med J, 1923. **1**(3236): p. 8-12.
3. Barnett, K., et al., *Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study*. Lancet, 2012. **380**(9836): p. 37-43.
4. Redelmeier, D.A., S.H. Tan, and G.L. Booth, *The treatment of unrelated disorders in patients with chronic medical diseases*. N Engl J Med, 1998. **338**(21): p. 1516-20.
5. World Health Organization, *Primary Health Care-Now more than ever*. 2008, New York: The World Health Report.
6. Ornstein, S.M., et al., *The prevalence of chronic diseases and multimorbidity in primary care practice: a PPRNet report*. J Am Board Fam Med, 2013. **26**(5): p. 518-24.
7. Brett, T., et al., *Multimorbidity in patients attending 2 Australian primary care practices*. Ann Fam Med, 2013. **11**(6): p. 535-42.
8. Rizza, A., et al., *Age- and gender-related prevalence of multimorbidity in primary care: the Swiss FIRE project*. BMC Fam Pract, 2012. **13**: p. 113.
9. Wolff, J.L., B. Starfield, and G. Anderson, *Prevalence, expenditures, and complications of multiple chronic conditions in the elderly*. Arch Intern Med, 2002. **162**(20): p. 2269-76.
10. Fortin, M., et al., *Prevalence estimates of multimorbidity: a comparative study of two sources*. BMC Health Serv Res, 2010. **10**: p. 111.
11. Harrison, C., et al., *Examining different measures of multimorbidity, using a large prospective cross-sectional study in Australian general practice*. BMJ Open, 2014. **4**(7): p. e004694.
12. Fortin, M., et al., *Lifestyle factors and multimorbidity: a cross sectional study*. BMC Public Health, 2014. **14**: p. 686.
13. van Weel, C. and F.G. Schellevis, *Comorbidity and guidelines: conflicting interests*. Lancet, 2006. **367**(9510): p. 550-1.
14. Prados-Torres, A., et al., *Multimorbidity patterns: a systematic review*. J Clin Epidemiol, 2014. **67**(3): p. 254-66.
15. Guthrie, B., et al., *Adapting clinical guidelines to take account of multimorbidity*. BMJ, 2012. **345**: p. e6341.
16. Herndon, M.B., et al., *Implications of expanding disease definitions: the case of osteoporosis*. Health Aff (Millwood), 2007. **26**(6): p. 1702-11.
17. Alwan, A., *Global status report on noncommunicable diseases 2010*. 2011, Geneva: World Health Organization.

18. Nuyen, J., et al., *Accuracy of diagnosing depression in primary care: the impact of chronic somatic and psychiatric co-morbidity*. Psychol Med, 2005. **35**(8): p. 1185-95.
19. Tinetti, M.E. and T. Fried, *The end of the disease era*. Am J Med, 2004. **116**(3): p. 179-85.
20. Hempstead, K., et al., *The fragmentation of hospital use among a cohort of high utilizers: implications for emerging care coordination strategies for patients with multiple chronic conditions*. Med Care, 2014. **52 Suppl 3**: p. S67-74.
21. Holman, H., *Chronic disease--the need for a new clinical education*. JAMA, 2004. **292**(9): p. 1057-9.
22. Cassel, C.K. and J.A. Guest, *Choosing wisely: helping physicians and patients make smart decisions about their care*. JAMA, 2012. **307**(17): p. 1801-2.
23. Smith, P.C., et al., *Missing clinical information during primary care visits*. JAMA, 2005. **293**(5): p. 565-71.
24. Kripalani, S., et al., *Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care*. JAMA, 2007. **297**(8): p. 831-41.
25. Wallace, E., et al., *Managing patients with multimorbidity in primary care*. BMJ, 2015. **350**(h176).
26. Haute Autorité de Santé, *Si vous souhaitez réaliser des recommandations de bonne pratique*. 2010.
27. Van Spall, H.G., et al., *Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review*. JAMA, 2007. **297**(11): p. 1233-40.
28. Gross, C.P., et al., *Reporting the recruitment process in clinical trials: who are these patients and how did they get there?* Ann Intern Med, 2002. **137**(1): p. 10-6.
29. Masoudi, F.A., et al., *Most hospitalized older persons do not meet the enrollment criteria for clinical trials in heart failure*. Am Heart J, 2003. **146**(2): p. 250-7.
30. Muth, C., et al., *Current guidelines poorly address multimorbidity: pilot of the interaction matrix method*. J Clin Epidemiol, 2014. **67**(11): p. 1242-50.
31. Boyd, C.M., et al., *Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance*. JAMA, 2005. **294**(6): p. 716-24.
32. Stone, N.J., et al., *2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines*. Circulation, 2014. **129**(25 Suppl 2): p. S1-45.
33. Montori, V.M., J.P. Brito, and H.H. Ting, *Patient-centered and practical application of new high cholesterol guidelines to prevent cardiovascular disease*. JAMA, 2014. **311**(5): p. 465-6.
34. Montori, V.M., J.P. Brito, and M.H. Murad, *The optimal practice of evidence-based medicine: incorporating patient preferences in practice guidelines*. JAMA, 2013. **310**(23): p. 2503-4.
35. Mol, A., *The logic of Care: health and the problem of patient choice*. 2008, London: Routledge.

36. Bury, M., *Chronic illness as biographical disruption*. *Sociol Health Illn*, 1982. **4**(2): p. 167-82.
37. Corbin, J. and A. Strauss, *Managing Chronic Illness at Home: Three lines of Work*. *Qualitative Sociology*, 1985. **8**(3): p. 224-246.
38. Russell, L.B., D.C. Suh, and M.A. Safford, *Time requirements for diabetes self-management: too much for many?* *J Fam Pract*, 2005. **54**(1): p. 52-6.
39. Jowsey, T., et al., *Time's up. descriptive epidemiology of multi-morbidity and time spent on health related activity by older Australians: a time use survey*. *PLoS One*, 2013. **8**(4): p. e59379.
40. Barber, N., et al., *Patients' problems with new medication for chronic conditions*. *Qual Saf Health Care*, 2004. **13**(3): p. 172-5.
41. Kesselheim, A.S., et al., *Burden of changes in pill appearance for patients receiving generic cardiovascular medications after myocardial infarction: cohort and nested case-control studies*. *Ann Intern Med*, 2014. **161**(2): p. 96-103.
42. Shippee, N.D., et al., *Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice*. *J Clin Epidemiol*, 2012. **65**(10): p. 1041-51.
43. Adams, C.D., et al., *Pediatric asthma: a look at adherence from the patient and family perspective*. *Curr Allergy Asthma Rep*, 2004. **4**(6): p. 425-32.
44. Osterberg, L. and T. Blaschke, *Adherence to medication*. *N Engl J Med*, 2005. **353**(5): p. 487-97.
45. Simpson, S.H., et al., *A meta-analysis of the association between adherence to drug therapy and mortality*. *BMJ*, 2006. **333**(7557): p. 15.
46. Ho, P.M., et al., *Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus*. *Arch Intern Med*, 2006. **166**(17): p. 1836-41.
47. Eton, D.T., et al., *Building a measurement framework of burden of treatment in complex patients with chronic conditions: a qualitative study*. *Patient Relat Outcome Meas*, 2012. **3**: p. 39-49.
48. Mair, F.S. and C.R. May, *Thinking about the burden of treatment*. *BMJ*, 2014. **349**: p. g6680.
49. Willis, J., *Foundations of Qualitative Research: Interpretive and Critical Approaches*. 2009, New York, NY: Sage.
50. Klopfenstein, C.E., et al., *Pain intensity and pain relief after surgery. A comparison between patients' reported assessments and nurses' and physicians' observations*. *Acta Anaesthesiol Scand*, 2000. **44**(1): p. 58-62.
51. Falissard, B., *Mesurer la subjectivité en santé. Perspective méthodologique et statistique*. 2eme édition ed. 2001, Paris: Masson. 211.
52. DeVellis, R.F., *Classical test theory*. *Med Care*, 2006. **44**(11 Suppl 3): p. S50-9.
53. DeVellis, R., *Scale Development: Theory and Application*, ed. Sage. 2003, Thousand Oaks, CA: Sage.
54. Streiner, D. and G. Norman, *Basic concepts*, in *Health Measurement Scales: a practical guide to their development and use*. 2008, Oxford University Press: Oxford. p. 6.

55. Wechsler, D., *The measurement of adult intelligence*. 1939, Baltimore: Williams & Wilkins.
56. Guyatt, G.H., et al., *Responsiveness and validity in health status measurement: a clarification*. J Clin Epidemiol, 1989. **42**(5): p. 403-8.
57. Scoggins, J.F. and D.L. Patrick, *The use of patient-reported outcomes instruments in registered clinical trials: evidence from ClinicalTrials.gov*. Contemp Clin Trials, 2009. **30**(4): p. 289-92.
58. Frank, L., E. Basch, and J.V. Selby, *The PCORI perspective on patient-centered outcomes research*. JAMA, 2014. **312**(15): p. 1513-4.
59. Marshall, M., et al., *Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia*. Br J Psychiatry, 2000. **176**: p. 249-52.
60. US Food and Drug Administration, *Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. 2009, Rockville: US Department of Health and Human Services.
61. Reeve, B.B., et al., *ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research*. Qual Life Res, 2013.
62. Terwee, C.B., et al., *Quality criteria were proposed for measurement properties of health status questionnaires*. J Clin Epidemiol, 2007. **60**(1): p. 34-42.
63. Shrout, P.E. and J.L. Fleiss, *Intraclass correlations: uses in assessing rater reliability*. Psychol Bull, 1979. **86**(2): p. 420-8.
64. Bland, J.M. and D.G. Altman, *Cronbach's alpha*. BMJ, 1997. **314**(7080): p. 572.
65. Eton, D.T., et al., *A systematic review of patient-reported measures of burden of treatment in three chronic diseases*. Patient Relat Outcome Meas, 2013. **4**: p. 7-20.
66. Guyatt, G.H., C. Bombardier, and P.X. Tugwell, *Measuring disease-specific quality of life in clinical trials*. CMAJ, 1986. **134**(8): p. 889-95.
67. Atkinson, M.J., et al., *Hierarchical construct validity of the treatment satisfaction questionnaire for medication (TSQM version II) among outpatient pharmacy consumers*. Value Health, 2005. **8 Suppl 1**: p. S9-S24.
68. Atkinson, M.J., et al., *Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease*. Health Qual Life Outcomes, 2004. **2**: p. 12.
69. Bland, J.M. and D.G. Altman, *Measuring agreement in method comparison studies*. Stat Methods Med Res, 1999. **8**(2): p. 135-60.
70. Gallacher, K., et al., *Uncovering treatment burden as a key concept for stroke care: a systematic review of qualitative research*. PLoS Med, 2013. **10**(6): p. e1001473.
71. Gallacher, K., et al., *Understanding patients' experiences of treatment burden in chronic heart failure using normalization process theory*. Ann Fam Med, 2011. **9**(3): p. 235-43.
72. Bohlen, K., et al., *Overwhelmed patients: a videographic analysis of how patients with type 2 diabetes and clinicians articulate and address treatment burden during clinical encounters*. Diabetes Care, 2012. **35**(1): p. 47-9.

73. Weiner, S.J., et al., *Contextual errors and failures in individualizing patient care: a multicenter study*. *Ann Intern Med*, 2010. **153**(2): p. 69-75.
74. Sperber, A.D., *Translation and validation of study instruments for cross-cultural research*. *Gastroenterology*, 2004. **126**(1 Suppl 1): p. S124-8.
75. Herdman, M., J. Fox-Rushby, and X. Badia, *A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach*. *Qual Life Res*, 1998. **7**(4): p. 323-35.
76. Herdman, M., J. Fox-Rushby, and X. Badia, *'Equivalence' and the translation and adaptation of health-related quality of life questionnaires*. *Qual Life Res*, 1997. **6**(3): p. 237-47.
77. Wild, D., et al., *Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation*. *Value Health*, 2005. **8**(2): p. 94-104.
78. Acquadro, C., et al., *Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials*. *Value Health*, 2008. **11**(3): p. 509-21.
79. Guillemin, F., C. Bombardier, and D. Beaton, *Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines*. *J Clin Epidemiol*, 1993. **46**(12): p. 1417-32.
80. Beaton, D.E., et al., *Guidelines for the process of cross-cultural adaptation of self-report measures*. *Spine (Phila Pa 1976)*, 2000. **25**(24): p. 3186-91.
81. Roberts, G., et al., *Enhancing rigour in the validation of patient reported outcome measures (PROMs): bridging linguistic and psychometric testing*. *Health Qual Life Outcomes*, 2012. **10**: p. 64.
82. Morisky, D.E., et al., *Predictive validity of a medication adherence measure in an outpatient setting*. *J Clin Hypertens (Greenwich)*, 2008. **10**(5): p. 348-54.
83. Morisky, D.E., L.W. Green, and D.M. Levine, *Concurrent and predictive validity of a self-reported measure of medication adherence*. *Med Care*, 1986. **24**(1): p. 67-74.
84. Slawsky, K., M. Massagli, and P. Wicks, *PRM30 a comparison of the PatientsLikeMe Quality of Life Questionnaire (PLMQOL) with the RAND SF-36*. *Value Health*, 2011. **14**(7): p. A426.
85. Bove, R., et al., *Evaluation of an online platform for multiple sclerosis research: patient description, validation of severity scale, and exploration of BMI effects on disease course*. *PLoS One*, 2013. **8**(3): p. e59707.
86. Wicks, P., K.A. Sulham, and A. Gnanasakthy, *Quality of life in organ transplant recipients participating in an online transplant community*. *Patient*, 2014. **7**(1): p. 73-84.
87. De la Loge, C., et al., *Characteristics of Users of the Epilepsy Community of PatientsLikeMe.com and Comparison with a Representative Claims Database*, in *63rd Annual Meeting of the American Academy of Neurology*. 2011: Honolulu, USA.
88. Ubel, P.A., A.P. Abernethy, and S.Y. Zafar, *Full disclosure--out-of-pocket costs as side effects*. *N Engl J Med*, 2013. **369**(16): p. 1484-6.

89. Lasch, K.E., et al., *PRO development: rigorous qualitative research as the crucial foundation*. Qual Life Res, 2010. **19**(8): p. 1087-96.
90. Patrick, D.L., et al., *Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument*. Value Health, 2011. **14**(8): p. 967-77.
91. May, C.R., et al., *Development of a theory of implementation and integration: Normalization Process Theory*. Implement Sci, 2009. **4**: p. 29.
92. Jani, B., et al., *Identifying treatment burden as an important concept for end of life care in those with advanced heart failure*. Curr Opin Support Palliat Care, 2013. **7**(1): p. 3-7.
93. Reinert, M., *Classification descendante hiérarchique: un algorithme pour le traitement des tableaux logiques de grandes dimensions(1986)Un logiciel d'analyse lexicale [ALCESTE]*. Cahiers de l'Analyse des Données, 1985: p. 471-484.
94. Reinert, M., *Une méthode de classification descendante hiérarchique: Application à l'analyse lexicale par contexte*. Cahiers de l'Analyse des Données, 1983. **3**: p. 187-198.
95. Ekman, A. and J.E. Litton, *New times, new needs; e-epidemiology*. Eur J Epidemiol, 2007. **22**(5): p. 285-92.
96. Glaser, B. and A. Strauss, *The discovery of grounded theory: Strategies for qualitative research*. 1967, Chicago, IL: Aldine Publishing Company.
97. Morse, J., *The Significance of Saturation*. Qual Health Res, 1995. **5**: p. 147-149.
98. Guest, G., A. Bunce, and L. Johnson, *How Many Interviews Are Enough? An Experiment with Data Saturation and Variability*. Field Methods, 2006. **18**(1): p. 59-82.
99. Cheek, J., *The Politics and Practices of Funding Qualitative Inquiry*, in *The SAGE Handbook of Qualitative Research*, N. Denzin and Y. Lincoln, Editors. 2011, SAGE: Thousand Oaks, CA.
100. Creswell, J., *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. 2002, Upper Saddle River, NJ: Pearson Education.
101. Sandelowski, M., *Sample size in qualitative research*. Res Nurs Health, 1995. **18**(2): p. 179-83.
102. Morse, J., *Designing funded qualitative research*, in *Handbook for Qualitative research*. 1994, Sage: Thousand Oaks, CA.
103. Baker, S. and R. Edwards, *How many qualitative interviews are enough? Expert voices and early career reflections on sampling and cases in qualitative research*, NCRM National Centre for Research methods, Editor. 2012.
104. Mason, M., *Sample size and Saturation in PhD Studies Using Qualitative Interviews*. Forum: Qualitative Social Research, 2010. **11**(3).
105. Francis, J.J., et al., *What is an adequate sample size? Operationalising data saturation for theory-based interview studies*. Psychol Health, 2010. **25**(10): p. 1229-45.
106. Marshall, M., *Sampling for qualitative research*. Family Practice, 1996. **13**(6): p. 522-525.

107. Califf, R.M., et al., *Characteristics of clinical trials registered in ClinicalTrials.gov, 2007-2010*. JAMA, 2012. **307**(17): p. 1838-47.
108. Duracinsky, M., et al., *The development of PROQOL-HIV: an international instrument to assess the health-related quality of life of persons living with HIV/AIDS*. J Acquir Immune Defic Syndr, 2012. **59**(5): p. 498-505.
109. Beresniak, A., et al., *Validation of a New International Quality-of-Life Instrument Specific to Cosmetics and Physical Appearance: BeautyQoL Questionnaire*. Arch Dermatol, 2012. **148**(11): p. 1275-82.
110. De Argaez, E. *Internet World Stats*. 2014 [cited 2014 13th january]; Available from: <http://www.internetworldstats.com/>.
111. Hewson, C., D. Laurent, and C. Vogel, *Proper methodologies for psychological and sociological studies conducted via the Internet*. Behavior Research Methods, Instruments, & Computers, 1996. **28**(2): p. 186-191.
112. Wicks, P., *Measuring the burden of treatment*. BMJ, 2014. **349**: p. g7322.
113. Kesse-Guyot, E., et al., *Participant profiles according to recruitment source in a large Web-based prospective study: experience from the Nutrinet-Sante study*. J Med Internet Res, 2013. **15**(9): p. e205.
114. Ekman, A., et al., *Feasibility of using web-based questionnaires in large population-based epidemiological studies*. Eur J Epidemiol, 2006. **21**(2): p. 103-11.
115. Kongsved, S.M., et al., *Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions*. J Med Internet Res, 2007. **9**(3): p. e25.
116. Sinnott, C., et al., *GPs' perspectives on the management of patients with multimorbidity: systematic review and synthesis of qualitative research*. BMJ Open, 2013. **3**(9): p. e003610.
117. Hughes, L.D., M.E. McMurdo, and B. Guthrie, *Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity*. Age Ageing, 2013. **42**(1): p. 62-9.
118. National Institute for Health and Care Excellence, *Scope consultation document, in Multimorbidity: clinical assessment and management*. 2014: Oxford.
119. Koch, H., et al., *Ordering blood tests for patients with unexplained fatigue in general practice: what does it yield? Results of the VAMPIRE trial*. Br J Gen Pract, 2009. **59**(561): p. e93-100.
120. American Geriatrics Society 2012 Beers Criteria Update Expert Panel, *American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults*. J Am Geriatr Soc, 2012. **60**(4): p. 616-31.
121. O'Mahony, D., et al., *STOPP/START criteria for potentially inappropriate prescribing in older people: version 2*. Age Ageing, 2014.
122. Gallagher, P. and D. O'Mahony, *STOPP (Screening Tool of Older Persons' potentially inappropriate Prescriptions): application to acutely ill elderly patients and comparison with Beers' criteria*. Age Ageing, 2008. **37**(6): p. 673-9.
123. Hamilton, H., et al., *Potentially inappropriate medications defined by STOPP criteria and the risk of adverse drug events in older hospitalized patients*. Arch Intern Med, 2011. **171**(11): p. 1013-9.

124. Weymiller, A.J., et al., *Helping patients with type 2 diabetes mellitus make treatment decisions: statin choice randomized trial*. Arch Intern Med, 2007. **167**(10): p. 1076-82.
125. Murad, M.H., et al., *Individuals with diabetes preferred that future trials use patient-important outcomes and provide pragmatic inferences*. J Clin Epidemiol, 2011. **64**(7): p. 743-8.
126. Stott, N.C. and R.H. Davis, *The exceptional potential in each primary care consultation*. J R Coll Gen Pract, 1979. **29**(201): p. 201-5.
127. Muth, C., et al., *The Ariadne principles: how to handle multimorbidity in primary care consultations*. BMC Med, 2014. **12**: p. 223.
128. Mechanic, D., D.D. McAlpine, and M. Rosenthal, *Are patients' office visits with physicians getting shorter?* N Engl J Med, 2001. **344**(3): p. 198-204.
129. Rothman, A.A. and E.H. Wagner, *Chronic illness management: what is the role of primary care?* Ann Intern Med, 2003. **138**(3): p. 256-61.
130. Ravaud, P., et al., *ARTIST (osteoarthritis intervention standardized) study of standardised consultation versus usual care for patients with osteoarthritis of the knee in primary care in France: pragmatic randomised controlled trial*. BMJ, 2009. **338**: p. b421.
131. Brilleman, S.L. and C. Salisbury, *Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study*. Fam Pract, 2013. **30**(2): p. 172-8.
132. Christensen, C., J. Grossman, and J. Hwang, *Disruptive solutions for the care of chronic disease*, in *The Innovator's prescription*. 2009, McGrawHill: New York.
133. Gittell, J.H., et al., *Impact of relational coordination on quality of care, postoperative pain and functioning, and length of stay: a nine-hospital study of surgical patients*. Med Care, 2000. **38**(8): p. 807-19.
134. Goroll, A.H. and D.P. Hunt, *Bridging the hospitalist-primary care divide through collaborative care*. N Engl J Med, 2015. **372**(4): p. 308-9.
135. Feltner, C., et al., *Transitional care interventions to prevent readmissions for persons with heart failure: a systematic review and meta-analysis*. Ann Intern Med, 2014. **160**(11): p. 774-84.
136. Fox, S. *The social life of health information*. 2011 [cited 2012 13/11]; Available from: http://www.pewinternet.org/~media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf
137. Wicks, P., et al., *Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm*. Nat Biotechnol, 2011. **29**(5): p. 411-4.
138. Lester, R.T., et al., *Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial*. Lancet, 2010. **376**(9755): p. 1838-45.
139. McManus, D.D., et al., *A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation*. Heart Rhythm, 2013. **10**(3): p. 315-9.

140. Hao, Y. and R. Foster, *Wireless body sensor networks for health-monitoring applications*. *Physiological Measurement*, 2008. **29**: p. 27-56.
141. Topol, E.J., S.R. Steinhubl, and A. Torkamani, *Digital medical tools and sensors*. *JAMA*, 2015. **313**(4): p. 353-4.
142. Atherton, H., et al., *Email for clinical communication between patients/caregivers and healthcare professionals*. *Cochrane Database Syst Rev*, 2012. **11**: p. CD007978.
143. Nami, N., et al., *Concordance and Time Estimation of Store-and-forward Mobile Teledermatology Compared to Classical Face-to-face Consultation*. *Acta Derm Venereol*, 2015. **95**(1): p. 35-9.
144. Rat, A.C., et al., *Content of quality-of-life instruments is affected by item-generation methods*. *Int J Qual Health Care*, 2007. **19**(6): p. 390-8.
145. Adelman, R.D., et al., *Caregiver burden: a clinical review*. *JAMA*, 2014. **311**(10): p. 1052-60.
146. Munjanja, S.P., G. Lindmark, and L. Nystrom, *Randomised controlled trial of a reduced-visits programme of antenatal care in Harare, Zimbabwe*. *Lancet*, 1996. **348**(9024): p. 364-9.
147. Emanuel, E.J. and V.R. Fuchs, *The perfect storm of overutilization*. *JAMA*, 2008. **299**(23): p. 2789-91.