

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO: CIÊNCIAS EM GASTROENTEROLOGIA
DISSERTAÇÃO DE MESTRADO**

**VALIDAÇÃO DE QUESTIONÁRIO DE MENSURAÇÃO DA DISPEPSIA
FUNCIONAL DESENVOLVIDO PARA A LÍNGUA PORTUGUESA**

GUILHERME BECKER SANDER

Dissertação apresentada ao Programa de Pós-Graduação: Ciências em Gastroenterologia como requisito parcial para obtenção do grau de mestre em Medicina: Gastroenterologia.

Orientador: Dr. José Roberto Goldim

**PORTO ALEGRE
2003**

S214v **Sander, Guilherme Becker**

Validação de questionário de mensuração da dispepsia funcional desenvolvido para a língua portuguesa / Guilherme Becker Sander; orient. José Roberto Goldim. Porto Alegre: UFRGS, 2003.

71p. gráf. tab.

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação: Ciências em Gastroenterologia.

1. DISPEPSIA. 2. QUESTIONÁRIOS. 3. SINAIS E SINTOMAS DIGESTÓRIOS. 4. REPRODUTIBILIDADE DE RESULTADOS. 5. ESTUDOS DE VALIDAÇÃO. 6. ANÁLISE FATORIAL. 7. PROJETOS DE PESQUISA. I. Goldim, José Roberto. II. Título

C.D.D. 616.332

C.D.U. 616.33-02(076.3)(043.3)=806.90

AGRADECIMENTOS

Ao Prof. Dr. Luiz Edmundo Mazzoleni, pelo modelo de médico e de pessoa que tem sido para mim e pela sua imensa bondade e fraternidade.

Ao Prof. Dr. José Roberto Goldim por ter me aceito como aluno, pela orientação desta tese e pelo apoio constante.

Aos Drs. André Castagna Wortmann, Andry Fiterman Costa e Eduardo André Ott, pelo auxílio, incentivo, amizade e coleguismo.

Aos colegas, Vincenzo Piccoli, Ângela Cristiane da Silva, Leandro Oliveira, Simone Beheregaray, Simone Matioti, Gustavo Somm pelo apoio na coleta dos dados deste estudo.

Aos meus pais, Sinay e Maria Aparecida, que tanto me orgulham pelos seus princípios de humanidade e dedicação aos filhos.

À minha noiva, Simone, pela compreensão, companheirismo, dedicação e amor.

Aos meus avós, Rachel, Dagoberto, Enaura e Ernesto, pela inspiração e luz que lançam no meu caminho.

Aos meus irmãos, Ernesto, Valéria, Luciana e Pedro, pelo carinho e pela torcida.

Ao Prof. Dr. Paulo Dornelles Picon pelo apoio, amizade, compreensão e incentivo.

Ao Prof. Dr. Carlos Fernando Francesconi pelo apoio e orientação.

Ao Programa de Pós-Graduação: Ciências em Gastroenterologia.

Aos professores do Programa de Pós-Graduação: Ciências em Gastroenterologia.

Aos professores do Serviço de Gastroenterologia do Hospital de Clínicas de Porto Alegre (HCPA).

Ao Grupo de Pesquisa e Pós-graduação do HCPA e ao Fundo de Incentivo à Pesquisa do HCPA pelo apoio financeiro na tradução deste trabalho para o inglês.

Aos médicos contratados Cristina, Helenice, Ilton e Mário, do serviço de Gastroenterologia do Hospital de Clínicas de Porto Alegre, pela orientação e amizade.

Aos amigos Fernando Soares, Sandra Schmaedecke e Terezinha Laggazio pelo apoio incondicional em todas as horas.

Aos pacientes e voluntários que participaram do estudo, pela sua vontade de ajudar ao próximo, disposição, tempo e amizade com a equipe.

SUMÁRIO

1. INTRODUÇÃO.....	5
2. REVISÃO DA LITERATURA.....	6
2.1. Definição	6
2.2. Critérios diagnósticos	6
2.3. Causas de dispepsia.....	6
2.4. Prevalência e impacto econômico	7
2.5. Mensuração da dispepsia.....	8
2.6. Desenvolvimento e validação de um questionário	12
REFERÊNCIAS BIBLIOGRÁFICAS	15
ANEXO I – VERSÃO PARA PUBLICAÇÃO EM PORTUGUÊS.....	18
Resumo	20
Introdução	21
Pacientes e Métodos	25
Resultados.....	30
Discussão	35
Conclusões.....	39
Referências bibliográficas	40
ANEXO II - QUESTIONÁRIO DE SINTOMAS DISPÉPTICOS.....	43
ANEXO III – VERSÃO PARA PUBLICAÇÃO EM INGLÊS.....	45
Abstract.....	47
Introduction.....	48
Patients and Methods	48
Results.....	57
Discussion.....	62
Conclusions.....	66
References.....	67
ANEXO IV – DYSPEPTIC SYMPTOMS QUESTIONNAIRE	70
ANEXO V - GRÁFICOS ADICIONAIS.....	72

1. INTRODUÇÃO

Define-se dispepsia por dor ou desconforto centrado no abdômen superior. Entre esses sintomas incluem-se dor, plenitude abdominal, saciedade precoce, náuseas, vômitos e sensação de distensão. Os critérios diagnósticos foram definidos por reuniões de consenso e são conhecidos por critérios de Roma.^{1,2} Apesar de muito prevalente, a dispepsia funcional permanece sendo uma entidade de difícil estudo pela falta de ferramentas adequadas para mensuração de desfechos significativos. Isto acontece porque a dispepsia funcional não possui um substrato anatômico ou fisiopatológico mensurável, e, portanto, torna-se obrigatória a valorização de aspectos subjetivos para se quantificar benefícios de intervenções terapêuticas. O objetivo deste estudo foi desenvolver e validar um questionário para avaliação de pacientes brasileiros com dispepsia funcional.

2. REVISÃO DA LITERATURA

2.1. Definição

Define-se dispepsia por dor ou desconforto centrado no abdômen superior. Entre esses sintomas incluem-se dor, plenitude abdominal, saciedade precoce, náuseas, vômitos e sensação de distensão. A definição de desconforto é variada, sendo que em algumas regiões do mundo significa uma dor leve, enquanto em outras é claramente distinguida da dor.¹ Estes sintomas não são obrigatoriamente relacionados com as refeições. Salienta-se que, com base em estudos de pHmetria de 24 hrs, pacientes em que o sintoma principal é a pirose quase sempre apresentam refluxo ácido patológico, devendo ser enquadrados como tendo doença do refluxo gastroesofágico e devem ser distinguidos dos pacientes com dispepsia.¹

2.2. Critérios diagnósticos

Os critérios diagnósticos para dispepsia funcional foram definidos em consensos de especialistas e são conhecidos como Critérios de Roma.^{1,2} Para o diagnóstico de dispepsia funcional por esses critérios são necessários a presença dos seguintes fatores, por pelo menos 12 semanas, não obrigatoriamente consecutivas, nos últimos 12 meses: (1) dispepsia persistente ou recorrente, (2) nenhuma evidência de doença orgânica que justifique os sintomas (necessária a realização de endoscopia digestiva alta), (3) sem evidência de que a dispepsia seja aliviada exclusivamente pela evacuação ou associada com alteração na frequência ou no formato das fezes.

2.3. Causas de dispepsia

Os pacientes dispépticos podem ser subdivididos em 3 grandes grupos: o primeiro é caracterizado pela presença de uma doença orgânica identificável em que exista uma relação de causalidade provada com sintomas de dispepsia (úlceras pépticas, câncer, uso de

medicamentos); um segundo com anormalidades fisiopatológicas ou microbiológicas presentes, mas cuja causalidade com a dispepsia não está bem definida (infecção pelo *Helicobacter pylori*, duodenite histológica, calculose biliar, hipersensibilidade visceral, dismotilidade gastroduodenal); e o terceiro sem nenhuma anormalidade identificável. Os pacientes do grupo 2 e 3, em que não há uma explicação definida para origem dos sintomas, formam o grupo de pacientes com dispepsia funcional.¹ Acredita-se que a dor na dispepsia funcional seja causada pela amplificação pelo sistema nervoso central de estímulos viscerais aferentes de modo que sejam percebidos conscientemente, por distúrbios da motilidade, ou por outros fatores mal definidos. Portanto, a dispepsia funcional não é uma doença orgânica nem psiquiátrica, mas parte de um quadro biopsicossocial relacionado a disfunção do eixo cérebro-intestino.³

Um outro tipo de dispepsia tem despertado grandes debates na literatura. Trata-se da “dispepsia não investigada”. Este grupo é formado por pacientes dispépticos cuja causa para os sintomas ainda não foi investigada. Vários algoritmos de manejo atualmente iniciam pelo tratamento empírico da dispepsia antes da investigação de sua causa,⁴ motivo pelo qual muitos autores têm utilizado esta expressão.

2.4. Prevalência e impacto econômico

Estudos populacionais têm mostrado prevalências de dispepsia que variam de 7% a 63% da população, sendo os achados tão diversos principalmente pela falta de consenso na definição de dispepsia e pelo tempo diferenciado de acompanhamento dos pacientes nos diferentes estudos. Considera-se, tendo como base a média obtida com os estudos já realizados, que a prevalência real esteja em torno de 25%.^{5,6} A prevalência da dispepsia é maior em mulheres e diminui com a idade. A dispepsia é responsável por 7% das consultas médicas de maneira geral e por 40 a 70% das queixas gastroenterológicas.⁶

O custo da dispepsia é tanto direto, incluindo atendimento médico, realização de exames e gastos com medicamentos, quanto indireto, representado pelo absenteísmo e diminuição da produtividade. O custo somente com medicamentos foi estimado, nos Estados Unidos em 1995, em 1,3 bilhão de dólares.⁶

Cabe ressaltar que a prevalência e custos da dispepsia foram estimados principalmente considerando a dispepsia não investigada, pois a maioria dos estudos foi realizada por levantamentos populacionais baseados em questionários de sintomas, sem estabelecimento da causa. Contudo, em menos de 40% dos pacientes dispépticos é encontrada causa identificável e cujo nexo causal com os sintomas pode ser estabelecido. Desta forma, dos dados de dispepsia não investigada, mais de 50% atribui-se à dispepsia funcional.⁶

Portanto, a dispepsia é um grande problema, tanto do ponto de vista clínico, quanto social e econômico.⁷ Tratamentos para a dispepsia funcional não têm demonstrado índices de sucesso acima de 40% sobre o placebo,⁸ sendo na maioria dos estudos menores que 20%, o que enfatiza a necessidade de mais estudos nesta área para o estabelecimento de condutas eficazes e custo-efetivas para o manejo desta situação.⁷

2.5. Mensuração da dispepsia

Apesar da grande prevalência e do custo sócio-econômico que acarreta, é a partir do final do século 20 que esta doença passou a ser mais estudada. Em parte pode-se atribuir este paradoxo às escassas ferramentas de pesquisa disponíveis. Em se tratando, por definição, de uma doença sem anormalidade estrutural ou fisiopatológica definida, os desfechos de estudos que avaliam eficácia terapêutica necessitam ser obrigatoriamente subjetivos. Torna-se então necessário o desenvolvimento de métodos que possibilitem que sintomas e percepções do paciente sejam mensurados, de maneira confiável e reprodutível, o que se tem conseguido, na maioria dos casos, através de questionários padronizados. As primeiras áreas do

conhecimento a se deparar com essas dificuldades foram a psicologia e a educação. Na medicina, diversas áreas têm se defrontado com esse dilema, que tem se tornado cada vez maior, uma vez que a ênfase na qualidade de vida tem se aproximado da ênfase na quantidade de vida.

Não existe na literatura um consenso de qual seria o melhor desfecho a ser utilizado em ensaios clínicos de dispepsia funcional. Várias escalas validadas de sintomas apresentam falhas metodológicas na sua construção ou validação. Em revisão sistemática sobre o tratamento da dispepsia funcional,⁹ dos 52 estudos incluídos, somente 5 foram realizados utilizando escalas validadas. As escalas criadas até então podem ser divididas em três categorias principais: escalas de avaliação global, instrumentos genéricos e escalas doença-específicas. Nas primeiras, uma impressão geral dos sintomas do paciente é questionada, sendo um exemplo a “*Seven Point Likert Scale*”, em que o paciente pode optar por uma graduação de 0 a 7 (variando de nenhum a muito acentuada) sobre a pergunta “*Overall, how would you rate the severity of your central upper abdominal pain/discomfort?*”.^{10,11} Os instrumentos genéricos são instrumentos de qualidade de vida que podem ser usados em diversas populações, sendo exemplos o questionário *short form-36*¹² e o instrumento de avaliação de qualidade de vida da Organização Mundial da Saúde, o WHOQOL.¹³ As escalas doença-específicas questionam o paciente sobre aspectos particulares da doença em questão, não podendo ser utilizados em outras situações clínicas. Podem ser unidimensionais, quando questionam basicamente sobre a sintomatologia apresentada, ou multidimensionais, quando incluem outros domínios, como por exemplo, o impacto sobre aspectos emocionais, sociais ou atividades cotidianas. Exemplos de escala unidimensional doença específica são a escala de Veldhuyzen van Zanten,¹⁴ a escala validada por Buckley¹⁵ e a “*Gastrointestinal Symptom Rating Scale (GSRS)*”.¹⁶ Como exemplo de escalas

multidimensionais temos a “*Glasgow Dyspepsia Severity Scale*”¹⁷ e a “*Nepean Dyspepsia Index*”.¹⁸

A escala validada por Veldhuyzen van Zanten, Tytgat¹⁴ foi montada pela seleção dos oito sintomas mais freqüentes em pacientes com dispepsia funcional. A seguir, foram aplicados os princípios da reprodutibilidade, responsividade e validade, concluindo que o questionário seria válido para avaliar sintomas dispépticos. Um dos problemas desta escala, é que os autores incluíram sintomas não classificados atualmente como dispépticos, tais como a pirose, “gosto azedo”, halitose e eructações. Uma segunda importante limitação deste estudo foi a realização de endoscopia digestiva alta entre a aplicação do primeiro e do segundo questionário, no período de avaliação da reprodutibilidade, sabendo-se que a endoscopia tem um forte efeito placebo, principalmente quando o resultado é normal.¹⁹ Em revisão recente,¹¹ o próprio Veldhuyzen van Zanten não cita sua escala como sendo de utilidade atualmente.

Outra referência da literatura na validação de questionário de sintomas dispépticos é uma publicação de Buckley e cols. de 1997,¹⁵ na qual os autores selecionaram quatro sintomas (dor epigástrica, pirose, eructações, distensão) considerados como os mais freqüentes em 50 dispépticos funcionais, e avaliaram esses sintomas com escores de 1 a 5 para intensidade, freqüência e duração dos sintomas, e concluíram que o questionário é um instrumento adequado para avaliar a dispepsia funcional. O questionário tem a grande vantagem de discriminar mais detalhadamente os sintomas, abordando não apenas a presença ou não dos mesmos, mas discriminando as suas múltiplas características (duração, intensidade, freqüência), e permitindo desta forma quantificar melhor a sintomatologia dispéptica funcional. Contudo, não inclui sintomas importantes na dispepsia como náuseas, e incluiu a pirose e a eructação como sintomas dispépticos, não estando de acordo com os consensos de Roma.^{1,2}

Outra escala muito utilizada em ensaios clínicos de dispepsia funcional é a “Gastrointestinal Symptom Rating Scale” (GSRS).¹⁶ A GSRS utiliza uma escala ordinal de sete pontos para mensurar vários sintomas, entre eles, dor abdominal, refluxo, indigestão, diarreia e constipação. Na verdade, a escala foi desenvolvida para ser usada em estudos de síndrome do intestino irritável e úlcera péptica, e não para dispepsia funcional.¹⁶ Um critério essencial para a criação de escalas aditivas é que elas realmente sejam unidimensionais, ou seja, meçam um único conceito,²⁰ o que parece não ser verdadeiro nesta escala.

Um dos desfechos mais freqüentemente utilizados nos grandes estudos, o “*Seven Point Likert Scale*”,¹⁰ por ser auto-aplicável, pode ser sujeito a diferentes interpretações pelos pacientes. Isto foi demonstrado pelo coeficiente de correlação intraclasse menor do que o esperado apresentado no estudo de validação da escala.¹⁰ Isto é especialmente verdadeiro se a escala é aplicada em locais em o que nível educacional da população não seja àqueles em que foi desenvolvida. Além disso, avalia sintomas apenas durante sete dias, período que se considera curto para uma patologia crônica e com severidade dos sintomas flutuantes.¹¹ Por medir somente dor e desconforto, pode causar um viés de seleção, visto que pacientes com dispepsia tipo dismotilidade tenderão a pontuar zero nessa escala, e provavelmente não serão incluídos nos estudos. Não obstante, atualmente também tem sido enfatizado o pequeno valor de comparação de médias entre grupos de tratamento, tendo sido dada preferência a resultados dicotômicos. Em escalas de avaliação global, o único desfecho dicotômico aceitável seria a proporção de pacientes com melhora completa dos sintomas. Melhoras parciais são difíceis de definir em escalas com poucos pontos, dada a dificuldade de estabelecer o quanto de melhora parcial seria clinicamente significativo em tais escalas. Contudo a literatura tem enfatizado que melhora completa é um objetivo irreal em dispepsia funcional.¹¹

A escala de Glasgow avalia principalmente as repercussões da dispepsia na vida dos pacientes, com pequena valorização dos sintomas dispépticos.¹⁷ Não obstante, o

questionário de Glasgow foi desenvolvido na Escócia em um ambiente de altíssima organização do sistema de saúde. Questões como número de visitas médicas domiciliares, número de consultas médicas nos últimos seis meses, uso de medicações não prescritas e uso de medicações prescritas muito provavelmente têm um efeito discriminante muito menor na nossa população. Além disso, a responsividade da escala foi determinada em pacientes com úlcera duodenal, não estando estabelecida para a dispepsia funcional.^{17,19} Outra crítica é a utilização de perguntas sobre os últimos 6 meses, pois não se sabe se os pacientes conseguem responder de forma acurada sobre eventos acontecidos em períodos tão longos.¹¹

A escala multidimensional “*Nepean Dyspepsia Index*” avalia a repercussão da dispepsia nas atividades de lazer, do cotidiano, do trabalho, assim como avalia o grau de temor e de ansiedade que os problemas digestivos determinam no paciente.¹⁸ É uma escala de auto-administração, incorrendo nos problemas citados anteriormente para esse tipo de avaliação.

2.6. Desenvolvimento e validação de um questionário

O processo de elaboração e de validação de um questionário precisa seguir etapas bem estabelecidas. Primeiramente deve-se escolher um grupo de sintomas que representem ao máximo a doença em questão.²¹ Deve ser construído de forma a obter o maior número de informações no menor número possível de perguntas, para que seja prático. Deve ter a maior capacidade possível de generalização para poder ser aplicado em diversas populações sem perder as suas características básicas.

Após a sua construção o questionário deve ser validado, ou seja, satisfazer quesitos básicos que comprovem a sua utilidade para pesquisa clínica. As propriedades de mensuração clínica, chamadas de propriedades clinimétricas, consideradas indispensáveis na validação de

um questionário de sintomas são a consistência interna, a reprodutibilidade, a responsividade e a validade.²²

Quando se cria um questionário para avaliação de uma situação clínica, especialmente quando se criam escalas aditivas, ou seja, que os resultados serão somados no final, espera-se que os diferentes itens desse questionário tenham, no mínimo, moderada relação entre si e o escore total, ou seja, espera-se que a escala seja homogênea, medindo diferentes aspectos da mesma condição, e não características diferentes de diferentes doenças, pois não haveria lógica em somá-las. Este atributo é conhecido por consistência interna e pode ser estimado pelo coeficiente de Cronbach,²³ indicado principalmente para escalas aditivas.²⁰ Valores muito elevados sugerem redundância nas questões, enquanto valores muito baixos sugerem que a escala esteja avaliando mais de um constructo.²⁴ Outra ferramenta utilizada para a avaliação das interrelações entre os itens é a análise fatorial. A análise fatorial tem tido um papel fundamental na determinação empírica da dimensionalidade dos itens. Um alto grau de correlação entre os itens sugere que todos os itens mensurem o mesmo conceito. Após esta análise, itens com baixa correlação com o escore total poderiam ser excluídos da escala.^{20,25}

A reprodutibilidade visa mostrar a fidedignidade dos dados obtidos, e pode ser aferida pelo método “teste-reteste”,²⁶ que é a obtenção de escores similares em duas ocasiões distintas, sem intervenções terapêuticas entre as duas aferições, geralmente realizadas em um intervalo suficientemente curto para não haver modificações na doença base¹⁸, e também que seja suficiente longo para que não haja viés de lembrança, ou seja, que o paciente responda na segunda ocasião pensando nas respostas que forneceu na primeira vez e não realmente aferindo os seus sintomas *de novo*.^{20,27}

A responsividade mostra a sensibilidade do questionário em demonstrar alterações quando eles realmente tenham acontecido.²⁸ Esta propriedade é testada com a aplicação do instrumento de aferição pré e pós-intervenção terapêutica que comprovadamente exerça

efeitos na doença, devendo então o questionário ter o poder de demonstrar as variações ocorridas. É uma característica essencial para instrumentos que serão utilizados em ensaios clínicos.

A validade se refere à capacidade de um instrumento aferir realmente o que se pretende que ele mensure.²² Os métodos empregados para investigar esta propriedade são numerosos e não há consenso sobre os nomes a serem utilizados para cada um dos métodos.²⁰ A validade de conteúdo não é medida por nenhum teste estatístico em particular, mas resulta do juízo de pessoas com reconhecido saber no assunto em julgar que os questionamentos propostos na escala são relevantes e representativos da situação se pretende medir. A validade de critério refere-se à capacidade de o instrumento de medida funcionar como preditor de outra variável, operacionalmente independente, mas que meça o mesmo constructo. Pode ser avaliada pela aplicação, paralelamente ao instrumento em estudo, de outro já cientificamente estabelecido como capaz de aferir aquela doença, devendo haver boa correlação entre os valores obtidos nos dois questionários.²⁰ Não havendo outro meio cientificamente estabelecido disponível que meça o mesmo constructo, pode-se utilizar medida de algum constructo relacionado ao que se pretende medir, esperando-se nesse caso coeficientes de correlação menores entre os instrumentos de medida, pela variabilidade imposta por estar-se medindo constructos, que apesar de relacionados, sejam diferentes.²⁹ Outra forma de validade, denominada por alguns de “validade discriminante”^{30,31} pode ser avaliada pela análise de grupos extremos, que é a capacidade de diferenciar uma população com de outra sem a doença através da comparação obtida com os escores da escala.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Talley N, Stanghellini V, Heading RC, Koch KL, Malagelada J, Tytgat G. Functional Gastrointestinal Disorders. In: Drossman DA, Corazziari E, Talley N, Thompson WG, Whitehead WE. The Functional Gastrointestinal Disorders. Second ed. Lawrence: Degnon Associates; 2000: 299-350.
2. Talley NJ, Colin-Jones D, Koch K, Koch M, Nyren O, Stanghellini V. Functional dyspepsia: a classification with guidelines for diagnosis and management. *Gastroenterol Int* 1991;4:145-160.
3. Drossman DA. Chronic Abdominal Pain (with emphasis on functional abdominal pain syndrome). In: Feldman M, Scharschmidt BF, Sleisenger MH. Feldman: Sleisenger & Fordtran's Gastrointestinal and Liver Disease. 7th ed. W.B. Saunders; 2002.
4. Jones RH. Approaches to uninvestigated dyspepsia. *Gut* 2002;50 Suppl 4:iv42-iv46.
5. Agreus L. Natural history of dyspepsia. *Gut* 2002;50 Suppl 4:iv2-iv9.
6. McQuaid KR. Dyspepsia. In: Feldman M, Scharschmidt BF, Sleisenger MH. Feldman: Sleisenger & Fordtran's Gastrointestinal and Liver Disease. 7th ed. W.B. Saunders; 2002.
7. Moayyedi P, Mason J. Clinical and economic consequences of dyspepsia in the community. *Gut* 2002;50 Suppl 4:iv10-iv12.
8. Bytzer P. H₂ receptor antagonists and prokinetics in dyspepsia: a critical review. *Gut* 2002;50 Suppl 4:iv58-iv62.
9. Veldhuyzen Van Zanten SJ, Cleary C, Talley NJ, Peterson TC, Nyren O, Bradley LA et al. Drug treatment of functional dyspepsia: a systematic analysis of trial methodology with recommendations for design of future trials. *Am J Gastroenterol* 1996;91(4):660-673.
10. Junghard O, Lauritsen K, Talley NJ, Wiklund IK. Validation of seven graded diary cards for severity of dyspeptic symptoms in patients with non ulcer dyspepsia. *Eur J Surg Suppl* 1998;(583):106-111.
11. Veldhuyzen Van Zanten SJ. Assessment of outcome in dyspepsia: has progress been made? *Gut* 2002;50 Suppl 4:iv23-iv25.
12. Stewart A, Hays R, Ware J. The MOS short-form general health survey. *Med Care* 1988;26:724-734.
13. Fleck M, Louzada S, Xavier M, Chachamovich E, Vieira RR, Santos L et al. Aplicação da versão em português do instrumento de avaliação de qualidade de vida d Organização Mundial da Saúde (WHOQOL-100). *Rev Saude Publica* 1999;33(2):198-205.
14. Veldhuyzen Van Zanten SJ, Tytgat KM, Pollak PT, Goldie J, Goodacre RL, Riddell RH et al. Can severity of symptoms be used as an outcome measure in trials of non-ulcer dyspepsia and *Helicobacter pylori* associated gastritis? *J Clin Epidemiol* 1993;46(3):273-279.

15. Buckley MJ, Scanlon C, McGurgan P, O'Morain CA. A validated dyspepsia symptom score. *Ital J Gastroenterol Hepatol* 1997;29(6):495-500.
16. Svedlund J, Sjödin I, Dotevall G. GSRS - a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. *Dig Dis Sci* 1988;33:129-134.
17. El Omar EM, Banerjee S, Wirz A, McColl KE. The Glasgow Dyspepsia Severity Score - a tool for the global measurement of dyspepsia. *Eur J Gastroenterol Hepatol* 1996;8(10):967-971.
18. Talley NJ, Verlinden M, Jones M. Quality of life in functional dyspepsia: responsiveness of the Nepean Dyspepsia Index and development of a new 10-item short form. *Aliment Pharmacol Ther* 2001;15(2):207-216.
19. Veldhuyzen Van Zanten SJ, Talley NJ, Blum AL, Bolling-Sternevald E, Sundin M, Junghard O. Combined analysis of the ORCHID and OCAY studies: does eradication of *Helicobacter pylori* lead to sustained improvement in functional dyspepsia symptoms? *Gut* 2002;50 Suppl 4:iv26-iv30.
20. Fachel J, Carney S. Avaliação psicométrica: a qualidade das medidas e o entendimento dos dados. In: Cunha J. *Psicodiagnóstico V*. Porto Alegre: Artmed Editora; 2000.
21. Kirschner B, Guyatt G. A Methodological Framework for Assessing Health Indices. *J Chron Dis* 1985;38(1):27-36.
22. Streiner DL, Norman GR. *Health Measurement Scales*. Second ed. Oxford: Oxford Medical Publications; 1995.
23. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16(3):297-334.
24. Selecting the items. In: Streiner DL, Norman GR. *Health Measurement Scales*. Second ed. Oxford: Oxford Medical Publications; 1995.
25. Shaw M, Talley NJ, Adlis S, Beebe T, Tomshine P, Healey M. Development of a digestive health status instrument: tests of scaling assumptions, structure and reliability in a primary care population. *Aliment Pharmacol Ther* 1998;12(11):1067-1078.
26. O'Leary MP, Barry MJ, Fowler FJ, Jr. Hard measures of subjective outcomes: validating symptom indexes in urology. *J Urol* 1992;148(5):1546-1548.
27. Reliability. In: Streiner DL, Norman GR. *Health Measurement Scales*. Second ed. Oxford: Oxford Medical Publications; 1995.
28. Liang MH. Evaluating measurement responsiveness. *J Rheumatol* 1995;22(6):1191-1192.
29. Validity. In: Streiner DL, Norman GR. *Health Measurement Scales*. Second ed. Oxford: Oxford Medical Publications; 1995.

30. Fleck MP, Louzada S, Xavier M, Chachamovich E, Vieira G, Santos L et al. [Application of the Portuguese version of the abbreviated instrument of quality life WHOQOL-bref]. Rev Saude Publica 2000;34(2):178-183.
31. Chassany O, Marquis P, Scherrer B, Read NW, Finger T, Bergmann JF et al. Validation of a specific quality of life questionnaire for functional digestive disorders. Gut 1999;44(4):527-533.

ANEXO I – VERSÃO PARA PUBLICAÇÃO EM PORTUGUÊS

Desenvolvimento e Validação de um Questionário em Português do Brasil para Avaliação da Dispepsia Funcional

Guilherme Becker Sander

Luiz Edmundo Mazzoleni

Carlos Fernando de Magalhães Francesconi

André Castagna Wortmann

Eduardo André Ott

Alexandro Theil

Vicenzo da Cruz Piccoli

Ângela Cristiane da Silva

Leandro Oliveira

Simone Beheregaray

Simoni Matioti

Gustavo Somm

José Roberto Goldim

Programa de Pós-graduação: Ciências em Gastroenterologia. Universidade Federal do Rio Grande do Sul. Hospital de Clínicas de Porto Alegre.

Endereço para correspondência:

Dr. Guilherme Becker Sander

Ramiro Barcelos, 1660 apto. 501. - CEP 90035-002 Porto Alegre. Brasil.

Telefone: +55-51-3316.8751 - Fax: +55-51-3316.8752

e-mail: guilhermesander@hotmail.com

Agradecimentos:

Nós agradecemos as seguintes pessoas pela sua contribuição para o desenvolvimento deste trabalho: Professor Sérgio Gabriel Barros, Dr. Júlio Pereira Lima, Dra Carisi Polanczyk, Dra Maria Helena Itaquí Lopes, Dr Cleber Kruehl.

Resumo

Introdução: Apesar de muito prevalente, a dispepsia funcional permanece sendo uma entidade de difícil estudo pela falta de ferramentas adequadas para mensuração de desfechos significativos. O objetivo deste estudo foi desenvolver e validar um questionário focado em sintomas, doença-específico para avaliação de pacientes brasileiros com dispepsia funcional. Para isso, o questionário foi cuidadosamente escrito seguindo terminologia amplamente aceita, de forma a facilitar a tradução e validação em outras línguas e culturas. O questionário foi desenvolvido utilizando a terminologia de Roma I para sintomas, que foram avaliados de acordo com a intensidade, duração e frequência quando possível. Trinta e um pacientes com dispepsia funcional, assim como 31 controles pareados por sexo e idade sem sintomas digestivos foram utilizados para avaliação da consistência interna, reprodutibilidade, responsividade, validade de conteúdo e validade discriminante do questionário. Outros 31 pacientes com dispepsia funcional foram incluídos para avaliação da validade de critério. O coeficiente alfa de Cronbach foi de 0,82. O coeficiente de correlação intraclassa para os escores obtidos com 7 dias de intervalo foi de 0,86. O escore médio obtido após 3 meses foi 16,4 vs. 23,03 na avaliação basal ($P=0,001$). Os dois gastroenterologistas “cegos” concordaram que o questionário avalia adequadamente a dispepsia funcional. A mediana dos escores dos controles foi 0 vs. 22,5 pontos nos pacientes dispépticos ($P=0,001$). Houve uma correlação inversa entre qualidade de vida e sintomas dispépticos ($R=-0,28;P=0,026$). O questionário proposto teve altos níveis de reprodutibilidade e de responsividade. Como o questionário foi baseado na terminologia utilizada no Consenso Internacional de Roma I, é esperado que ele será facilmente traduzido e validado.

Descritores: Dyspepsia, Questionnaires, Questionnaire Design, Reproducibility of Results

Introdução

A dispepsia é um problema médico muito prevalente que acomete cerca de 25% da população, estando a incidência anual em torno de 1%. Define-se dispepsia por dor ou desconforto centrado no abdômen superior.¹ Entre esses sintomas incluem-se dor, plenitude abdominal, saciedade precoce, náuseas, vômitos e sensação de distensão. Muitos indivíduos não apresentam alterações estruturais ou bioquímicas que expliquem os seus sintomas, e, portanto, são classificados como tendo dispepsia não ulcerosa ou funcional.¹ A fisiopatologia da dispepsia funcional é pouco compreendida, ela pode ser associada com alterações na motilidade, secreção e sensibilidade do trato digestório.²⁻⁴

Os critérios diagnósticos para dispepsia funcional foram definidos em consensos de especialistas e são conhecidos como Critérios de Roma.^{1,5} De acordo com esses critérios, a dispepsia funcional é diagnosticada se os seguintes sintomas estão presentes por pelo menos 12 semanas (não obrigatoriamente consecutivas) nos últimos 12 meses: (1) dispepsia persistente ou recorrente, (2) nenhuma evidência de doença orgânica que justifique os sintomas (incluindo a realização de endoscopia digestiva alta), (3) sem evidência de que a dispepsia seja aliviada exclusivamente pela evacuação ou associada com alteração na frequência ou no formato das fezes.

Apesar da grande prevalência e custo sócio-econômico que acarreta a partir do final do século 20 é que esta doença passou a ser mais estudada. Em parte pode-se atribuir este paradoxo às escassas ferramentas de pesquisa disponíveis. Em se tratando, por definição, de uma doença sem anormalidade estrutural ou fisiopatológica definida, os desfechos de estudos que avaliem eficácia terapêutica necessitam ser obrigatoriamente subjetivos. Torna-se então necessário o desenvolvimento de métodos que possibilitem que sintomas e percepções do paciente sejam mensurados, de maneira confiável e reprodutível. Isto tem sido conseguido através de questionários padronizados.

O processo de elaboração e de validação de um questionário precisa seguir etapas bem estabelecidas. Primeiramente deve-se escolher um grupo de sintomas que representem o máximo a doença em questão.⁶ Deve ser construído de forma a obter o maior número de informações no menor número possível de perguntas, para que seja prático. Deve ter a maior capacidade possível de generalização, para poder ser aplicado em diversas populações sem perder as suas características básicas. Deve-se atentar também para a possibilidade de futuras traduções, já que cada vez mais um número maior de estudos multicêntricos são realizados. O desafio não é criar um questionário fácil de ser traduzido, mas sim, garantir a comparabilidade dos resultados entre populações com diferenças étnicas e culturais.⁷

Após a sua construção o questionário deve ser validado, ou seja, satisfazer quesitos básicos que comprovem a sua utilidade para pesquisa clínica. As propriedades de mensuração clínica, chamadas de propriedades clinimétricas, consideradas indispensáveis na validação de um questionário de sintomas são a consistência interna, a reprodutibilidade, a responsividade e a validade.⁸

Espera-se que os diferentes itens deste questionário tenham, no mínimo, moderada relação entre si e entre o escore total, ou seja, esperamos que a escala seja homogênea, pois esperamos que a escala meça diferentes aspectos da mesma condição e não características diferentes de diferentes doenças. Este atributo é conhecido por consistência interna e pode ser estimado pelo coeficiente de Cronbach⁹. Valores muito elevados sugerem redundância nas questões, enquanto valores muito baixos sugerem que a escala esteja avaliando mais de um constructo, sendo o ideal que os valores estejam entre 0,7 e 0,9.¹⁰ Outra ferramenta utilizada para a avaliação das interrelações entre os itens é a análise fatorial. A análise fatorial tem tido um papel fundamental na determinação empírica da dimensionalidade dos itens. Um alto grau de correlação entre os itens sugere que todos os itens mensurem o mesmo conceito. Após esta análise, itens com baixa correlação com o escore total poderiam ser excluídos da escala.^{11,12}

A reprodutibilidade visa mostrar a fidedignidade dos dados obtidos, e pode ser aferida pelo método “teste-reteste”,¹³ que é a obtenção de escores similares em duas ocasiões distintas sem intervenções terapêuticas entre as duas aferições. As avaliações devem ser realizadas em um intervalo suficientemente curto para não haver modificações na doença base,¹⁴ mas também que seja suficiente longo para que não haja viés de lembrança, ou seja, que o paciente responda na segunda ocasião pensando nas respostas que forneceu na primeira vez e não realmente aferindo os seus sintomas *de novo*.¹⁵

A responsividade mostra a sensibilidade do questionário em demonstrar alterações quando eles realmente tenham acontecido.¹⁶ Esta propriedade é testada com a aplicação do instrumento de aferição pré e pós-intervenção terapêutica que comprovadamente exerça efeitos na doença.

A validade se refere a capacidade de um instrumento aferir realmente o que se pretende que ele mensure.⁸ Pode ser avaliada de várias maneiras. A validade de conteúdo não é uma análise estatística como as outras análises, mas um julgamento por especialistas no assunto sobre a representatividade e relevância dos itens propostos na escala.¹¹ Outra forma de validade é a análise de grupos extremos, que é a capacidade de diferenciar uma população com de outra sem a doença. Este tipo de validade também é denominada de “validade discriminante”.^{17,18} Também pode ser avaliada pela “validade de critério” em que paralelamente ao instrumento em estudo aplica-se outro já cientificamente estabelecido como capaz de aferir aquela doença, devendo haver boa correlação entre os valores obtidos nos dois questionários. Não havendo outro meio cientificamente estabelecido disponível, pode-se utilizar medida de algum constructo relacionado, esperando-se nesse caso coeficientes de correlação menores.¹⁸⁻¹⁹

O objetivo de nosso estudo foi o de desenvolver e validar um questionário fácil com facilidade de tradução, focado em sintomas (unidimensional) para a avaliação de

dispepsia funcional. Nós escolhemos um questionário unidimensional neste contexto por duas razões. Primeiro, questões de qualidade de vida são mais susceptíveis a erros de adaptação cultural durante futuros processos de tradução. Segundo, a atividade da dispepsia não ulcerosa é fortemente associada com a qualidade de vida. The fato, é frequentemente difícil determinar a direção da relação de causa e efeito destas entidades, o que pode fragilizar o uso de desfechos baseados em qualidade de vida na avaliação da dispepsia funcional. Desta forma, nós desenvolvemos um questionário em inglês baseado nos sintomas descritos como característicos de dipepsia no Consenso de Roma I. Para testar o potencial de tradução e as propriedades clinimétricas do questionário, ele foi traduzido para o português do Brasil e validado em um grupo de pacientes brasileiros com dispepsia funcional.

Métodos

Desenvolvimento

O questionário proposto no presente estudo, “*PADYQ – Porto Alegre Dyspeptic Symptoms Questionnaire*”, foi desenvolvido com base nos sintomas caracterizados como relevantes na reunião de especialistas de Roma I,⁵ resultando em um questionário com 11 itens (tabela 1). Todos os questionários foram respondidos com a assistência de um entrevistador. Os pacientes foram perguntados sobre a média dos seus sintomas durante o mês precedente.

Tabela 1 – Sintomas característicos de dispepsia funcional, avaliados pelo *Porto Alegre Dyspeptic Symptoms Questionnaire*

Sintoma	Pontuação
Dor no andar superior do abdômen	
Intensidade	0-5
Duração	0-3
Frequência	0-4
Náuseas	
Intensidade	0-5
Duração	0-3
Frequência	0-4
Vômitos	
Frequência	0-4
Sensação de distensão abdominal superior	
Intensidade	0-5
Duração	0-3

	Frequência	0-4
Saciedade precoce		
	Frequência	0-4
Total		0-44

Pacientes

Análise fatorial foi realizado levando em consideração todos os questionários respondidos pelos pacientes de um estudo sobre os efeitos da erradicação do *Helicobacter pylori* nos sintomas da dispepsia funcional (n=157). Todos os pacientes foram selecionados de acordo com os critérios de Roma I.

Para avaliação da consistência interna, reprodutibilidade e responsividade, 31 pacientes com sintomas dispépticos segundo critérios de Roma I foram incluídos no estudo (Grupo A). Todos realizaram endoscopia digestiva alta no Hospital de Clínicas de Porto Alegre com resultado normal.

Para a avaliação da validade discriminante, 31 pacientes pareados por sexo e idade do grupo de pacientes descrito acima e que relataram não ter problemas digestivos foram selecionados e entrevistados após a assinatura de consentimento informado (grupo controle).

Para a avaliação de validade de critério, o grupo controle e um grupo adicional de 31 pacientes dispépticos funcionais de acordo com os critérios de Roma I (Grupo B) responderam ao *Porto Alegre Dyspeptic Symptoms Questionnaire* e ao *World Health Organization Quality of Life (WHOQOL-BREF) questionnaire*.²⁰

O protocolo foi aprovado pelo Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre e todos os pacientes e controles assinaram termo de consentimento informado.

Avaliação das Propriedades Clinimétricas

As propriedades clinimétricas do Questionário de Sintomas Dispépticos foram avaliadas de 7 maneiras diferentes, apresentadas a seguir.

Análise fatorial. Análise fatorial utilizando o método de componentes principais foi utilizada para agregar itens em sub-escalas (domínios). Foram selecionados componentes com valores de raiz latente (“*eigenvalues*”) maiores que um. Após a extração dos componentes, foi realizada rotação ortogonal dos fatores pelo método Varimax de Kaiser. Foram agregados em domínios itens com carga fatorial maior que 0,4 no mesmo componente e que não tivessem valores maiores que 0,4 em outros componentes.²¹

Consistência interna. Consistência interna foi calculada com base nos questionários basais (T0) do Grupo A, sendo determinada pelo coeficiente de Cronbach.⁹ O resultado esperado para este coeficiente situa-se entre 0,7 e 0,9, que significaria redundância.^{10,12}

Reprodutibilidade. Os pacientes incluídos no estudo (Grupo A) responderam ao PADYQ na primeira consulta (consulta T0). Essa consulta foi repetida 7 dias após (consulta T1), quando os pacientes responderam novamente o questionário. Os pacientes não foram expostos a nenhum tipo de intervenção, diagnóstica ou terapêutica, entre T0 e T1. Em 15 pacientes o mesmo entrevistador preencheu o questionário na consulta T0 e T1 e em 16 pacientes o entrevistador nas duas consultas foi diferente, a fim de avaliar-se a concordância intra e interobservador. A correlação entre os escores foi estabelecida através do coeficiente de correlação intraclassa, calculado para os escores totais e para os diferentes domínios estabelecidos pela análise fatorial.

Responsividade. Após estas duas avaliações iniciais (consultas T0 e T1), os pacientes do Grupo A foram randomizados para um de dois grupos: um grupo recebeu 10 dias de lansoprazol, 30 mg, claritromicina, 500 mg e amoxicilina, 1g, todos duas vezes por dia. O

outro grupo recebeu lansoprazol, 30 mg duas vezes por dia, associados a placebos idênticos em aparência aos antibióticos utilizados pelo outro grupo, e com a mesma posologia. Estes grupos foram considerados como sendo homogêneos, pois não há evidência inequívoca na literatura de que haja de que um tratamento seja mais eficaz do que o outro.²² Uma diminuição de 13% a 73% nos escores foi esperada, de acordo com a resposta placebo obtida em estudos anteriores.²³ Após três meses do início do tratamento, uma nova consulta foi realizada (T2) e o PADYQ foi repetido por um dos entrevistadores. A responsividade foi calculada utilizando-se teste *t* para amostras pareadas, comparando-se o valor de T1 com T2 do escore total e dos diferentes domínios avaliados pelo questionário.

Validade de conteúdo. A validade de conteúdo foi avaliada por dois gastroenterologistas experientes “cegos” para o objetivo do questionário. Eles foram solicitados a descobrir qual constructo o questionário pretendia medir. Após, eles também foram solicitados a analisar se o questionário avaliava de maneira completa os sintomas dispépticos.

Validade discriminante. Os escores da consulta T0 desses pacientes dispépticos foram comparados com os escores do grupo controle. Os valores de casos e controles foram utilizando-se o teste de Wilcoxon.

Validade de critério. O PADYQ foi aplicado simultaneamente com o WHOQOL-BREF²⁰ (Questionário de Qualidade de Vida da Organização Mundial da Saúde) aos 31 pacientes controles e a 31 pacientes dispépticos (Grupo B). Os resultados foram correlacionados através do coeficiente de correlação não paramétrico de Spearman.

Tamanho da amostra. O tamanho da amostra necessária para o teste de reprodutibilidade foi calculado estimando uma correlação intraclasse de 0,9 (e considerando um erro alfa de 0,05 e aceitando um intervalo de confiança de $\pm 0,10$). De acordo com Streiner²⁴, o número de pacientes necessários para este propósito é de 31. O tamanho da

amostra para a avaliação de responsividade foi calculada para permitir a comparação pelo teste *t* de amostras pareadas, considerando-se um erro alfa de 0,05, um poder de 90%, uma redução esperada nos escores de 20% sobre escores basais com média de 25, um desvio padrão de 8 em ambas avaliações, com um coeficiente de correlação de 0,5 entre as visitas. O tamanho da amostra para a responsividade assim estimado foi de 27 pacientes. Para análise fatorial, sugere-se analisar pelo menos 5 vezes mais questionários do que o número de itens na escala, sendo considerado um mínimo de 100 questionários analisados para ter-se respostas adequadas.²¹ Existe dificuldade de cálculo do tamanho de amostra para outras propriedades clinimétricas, tendo sido utilizado da mesma forma os 30 pacientes estimados para a reprodutibilidade e responsividade, estando de acordo com outros estudos de validação de questionários dispépticos que usaram amostras semelhantes.^{25,26}

Resultados

Características da amostra

Dos 31 pacientes selecionados, 24 eram mulheres. O pareamento por sexo no grupo controle foi perfeito. A idade média dos pacientes com dispepsia foi de 43,1 anos (Grupo A), e a idade média dos controles foi de 43,0 (Coeficiente de correlação de Pearson $>0,99$). A escolaridade foi superior no grupo controle do que nos pacientes com dispepsia: 67% dos pacientes no grupo controle tinham nível superior, enquanto 58% dos pacientes dispépticos tinham menos do que 8 anos de escolaridade.

Análise fatorial

A análise fatorial dos onze itens dos questionários identificou 3 componentes com valores de raiz latentes maiores que 1 (figura 1). O primeiro componente, com valor de raiz latente de 3,077 e responsável por 24% da variância total do questionário, incluiu a intensidade, a duração e a frequência da distensão, assim como a frequência de saciedade precoce (cargas fatoriais após rotação Varimax de 0,937; 0,876; 0,910 e 0,433 respectivamente). O segundo componente, com valor de raiz latente de 2,4 e responsável por 24,3% da variância, incluiu intensidade, duração e frequência da náusea e frequência de vômitos (cargas fatoriais após rotação Varimax de 0,912; 0,847; 0,873; 0,577). O terceiro componente, com valor de raiz latente de 2,224 e responsável por 21,1% da variância, incluiu intensidade, duração e frequência da dor abdominal superior (cargas fatoriais após rotação Varimax de 0,905; 0,841; 0,867). Nenhum item teve carga fatorial acima de 0,2 em dois componentes simultaneamente (tabela 2 e Anexo 5-figura 4). Dessa forma denominamos o primeiro componente de domínio Distensão/Saciedade, o segundo de domínio Náuseas/Vômitos e o terceiro de domínio Dor. A variância somada dos três domínios explicou 70 % da variância do questionário.

Tabela 2 – Cargas fatoriais dos itens do Questionário de Sintomas Dispépticos após rotação Varimax de Kaiser

	Componente		
	1	2	3
Intensidade da dor	0,090	0,032	0,905
Duração da dor	-0,036	0,070	0,841
Frequência da dor	0,052	0,027	0,867
Intensidade das náuseas	0,034	0,912	0,003
Duração das náuseas	-0,009	0,847	-0,025
Frequência das náuseas	0,072	0,873	0,021
Frequência dos vômitos	0,142	0,577	0,141
Intensidade da distensão	0,937	0,038	0,037
Duração da distensão	0,876	0,004	-0,011
Frequência da distensão	0,910	0,034	-0,074
Frequência da saciedade precoce	0,433	0,155	0,144

Consistência interna

O coeficiente alfa de Cronbach para os 11 itens do *PADYQ* respondidos pelos 31 pacientes com dispepsia funcional (Grupo A) durante a consulta T1 foi de 0,82. O coeficiente alfa de Cronbach para o domínio Distensão/Saciedade foi de 0,75; o do domínio Náuseas/Vômitos 0,90 e o do domínio Dor 0,83.

Reprodutibilidade

O coeficiente de correlação intraclasse quando aplicado entre os escores totais das consultas T0 e T1 foi de 0,86 (figura 1). O valor obtido quando as entrevistas foram realizadas pelo mesmo entrevistador foi de 0,86 e com entrevistador diferente 0,87. O coeficiente de correlação intraclasse para o domínio Distensão/Saciedade foi de 0,87; o do domínio Náuseas/Vômitos 0,88 e o do domínio Dor 0,68. A média dos escores totais foi de 22,2 (DP 1,56) nas consultas T0 e de 23,2 (DP 1,42) nas consultas T1 ($P > 0,2$).

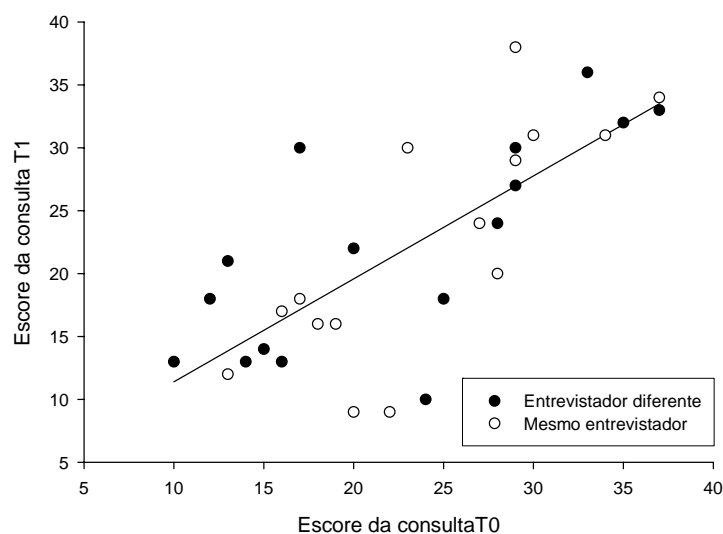


Figura 1. Correlação entre os escores obtidos pela aplicação do questionário nas visitas T0 e T1

Responsividade

A média dos escores na consulta T1 (pré-tratamento) foi de 23,2. Após tratamento medicamentoso, na consulta de avaliação com 3 meses de seguimento, a média dos escores era de 16,0 pontos ($P < 0,001$). Nos pacientes que receberam lansoprazol e placebos ($n=10$), a média dos escores era de 18,8 pontos (*versus* um valor basal de 25,9) ($P=0,038$). Nos pacientes que foram tratados com antibióticos ($n=21$), a média dos escores após 3 meses era de 14,67 pontos (contra um valor basal de 20,48) ($P=0,015$). O valor pré-tratamento do domínio Distensão/Saciedade foi de 9,5 *versus* 6,9 pontos 3 meses após ($P < 0,001$); no domínio Náuseas/Vômitos 5,36 *versus* 3,6 ($P=0,082$) e no domínio Dor 7,4 *versus* 5,5 ($P=0,016$).

Validade de conteúdo

Os dois gastroenterologistas “cegos” identificaram que o questionário avaliava dispepsia funcional. Eles consideram que os itens avaliam de maneira abrangente o espectro

da doença e que todos os itens eram relevantes. A clareza das questões também foi considerada adequada.

Validade discriminante

A média dos escores dos controles foi de 0,9 pontos e a mediana 0, tendo havido um mínimo de zero e um máximo de 7 pontos (figura 2). A média dos valores dos casos (Grupo A) foi de 22,2 pontos e a mediana 21,0 (mínimo = 9; máximo = 38). A diferença encontrada foi estatisticamente significativa ($P < 0,001$). Diferenças estatísticas foram encontradas nas comparações de cada um dos três domínios do questionário de casos e controles, com um $P < 0,01$ em todas três análises.

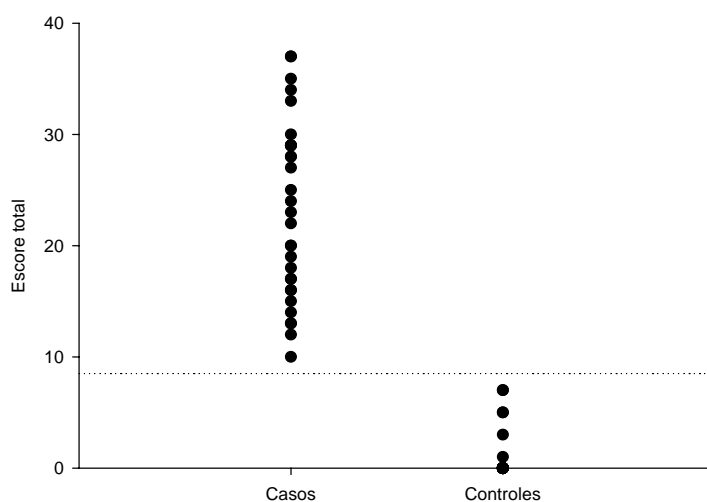


Figura 2 – Pontuação total de casos e controles

Validade de critério

Houve uma correlação estatisticamente significativa, embora fraca, entre os valores do PADYQ e o WHOQOL-BREF ($R = -0,288$; $P = 0,023$) para os 31 pacientes dispepsia do grupo B e os 31 controles. A correlação foi negativa, pois quanto mais alta a pontuação no PADYQ menor foi a do WHOQOL-BREF. Entre os domínios do PADYQ, o único que

mostrou correlação estatisticamente significativa com o questionário de qualidade de vida foi o domínio Distensão/Saciedade ($R=-0,352$; $P=0,005$).

Discussão

O questionário mostrou excelentes propriedades clinimétricas. O valor alfa de Cronbach, que mede a consistência interna (ou seja, o quanto cada item relaciona-se com os outros itens e com o escore total), mostrou-se dentro dos valores considerados com ideais (0,7 a 0,9)¹⁵, tanto quando o questionário de 11 itens foi analisado como um todo, quanto quando cada domínio foi analisado separadamente. O *Porto Alegre Dyspeptic Symptoms Questionnaire* (PADYQ) mostrou-se também reprodutível quando submetido ao procedimento de teste-reteste, tanto na avaliação com o mesmo entrevistador quanto na avaliação por entrevistadores diferentes. A responsividade também foi adequada a ensaios clínicos, já que em grupos de apenas 10 pacientes o questionário foi capaz de detectar alterações sintomáticas resultantes do efeito placebo. O PAQYQ foi também capaz de mostrar que mede o constructo correto, já que ele foi capaz de distinguir um grupo de pacientes sem doenças digestivas de um grupo de pacientes com sintomas crônicos no abdômen superior. A validade de conteúdo, avaliada por entrevistas qualitativas, mostrou que os itens do questionário são claros e relevantes.

A dispepsia é um problema muito freqüente na população e o seu tratamento ainda não é plenamente eficaz.²⁶ Uma das dificuldades que impedem o avanço do conhecimento neste campo é a dificuldade em mensurar com fidedignidade os benefícios alcançados ao propor-se um tratamento. O consenso de Roma II enfatiza a necessidade da utilização em ensaios clínicos de instrumentos de avaliação validados quando da mensuração de distúrbios funcionais do trato digestivo.¹ A aplicação de questionários desenvolvidos e validados em países desenvolvidos, com diferente perfil sócio-econômico e cultural, tem discutível aplicação em países em desenvolvimento, especialmente se considerarmos as questões de qualidade de vida e medidas de acesso aos serviços de saúde.

Alguns dos questionários existentes, na nossa opinião, apresentam problemas metodológicos. Por exemplo, o questionário validado por Valdhuyzen et al.²⁵ avalia sintomas não classificados atualmente como dispépticos, tais como a pirose, “gosto azedo”, halitose e eructações. Uma segunda importante limitação do processo de validação foi a realização de endoscopia digestiva alta entre a aplicação do primeiro e do segundo questionário, no período de avaliação da reprodutibilidade. O forte efeito placebo da endoscopia digestiva é bem conhecido, especialmente quando o resultado é normal.²²

O questionário validado por Buckley e cols. em 1997²⁶ tem igualmente problema na escolha dos sintomas representativos de dispepsia, tendo sido incluído no questionário pirose e eructações, e não avaliando outros sintomas dispépticos importantes como náuseas. Contudo, o questionário tem a grande vantagem de discriminar mais detalhadamente os sintomas, abordando não apenas a presença ou não dos mesmos, mas discriminando as suas múltiplas características (duração, intensidade, frequência), e permitindo desta forma quantificar melhor a sintomatologia dispéptica funcional.

Outra escala muito utilizada em ensaios clínicos de dispepsia funcional é a “Gastrointestinal Symptom Rating Scale” (GSRS). A GSRS utiliza uma escala ordinal de sete pontos para mensurar uma ampla gama de sintomas, muitos não considerados como parte da dispepsia funcional, como dor abdominal, refluxo, indigestão, diarreia e constipação. Na verdade, a escala foi desenvolvida para ser usada em estudos de síndrome do intestino irritável e úlcera péptica, a não para dispepsia funcional.³⁰

Um dos desfechos mais frequentemente utilizado nos grandes estudos de dispepsia funcional é o “Seven Point Likert Scale”.²⁹ É auto-aplicável e baseado em uma única pergunta: “Please state each day if you have experienced pain or discomfort in the stomach.” Dessa forma, pode ser sujeito a diferentes interpretações pelos pacientes, refletido no coeficiente de correlação intraclasse menor do que o esperado apresentado no estudo de

validação da escala.^{29,31} Isto é especialmente verdadeiro se a escala é aplicada em locais em o que nível educacional da população não seja àqueles em que foi desenvolvida. Além disso, avalia sintomas apenas durante sete dias, período que se considera curto para uma patologia crônica e com severidade dos sintomas flutuantes.²³ Finalmente, como os escores variam somente de 1 a 7, o poder discriminante é muito limitado.³¹

Outros questionários (“Critérios de Glasgow”, por exemplo), avaliaram principalmente as repercussões da dispepsia na vida dos pacientes, com pequena valorização dos sintomas dispépticos.²⁸ O questionário de Glasgow foi desenvolvido na Escócia em um ambiente de altíssima organização do sistema de saúde e inclui questões sobre o número de visitas médicas domiciliares, número de consultas médicas nos últimos seis meses, uso de medicações não prescritas e uso de medicações prescritas, que provavelmente têm um efeito discriminante muito menor em outras populações. Outro problema foi que a responsividade do questionário de Glasgow foi determinada em pacientes com úlcera duodenal, e, portanto, no nosso conhecimento, a validação do questionário de Glasgow não está plenamente determinada para pacientes com dispepsia funcional.²³

Quando o PAQYQ foi criado, o consenso de Roma II ainda não havia sido realizado, de modo que o questionário foi baseado nos critérios de Roma I. De fato, o consenso de Roma II introduziu poucas mudanças no que diz respeito a critérios diagnósticos e aos sintomas considerados como sendo de dispepsia funcional. Uma exceção importante é o vômito, presente entre os sintomas principais de Roma I, mas não nos de Roma II. Contudo, foi encontrada neste estudo uma forte correlação entre náuseas e vômitos em pacientes dispépticos funcionais, corroborando a idéia de que vômito deva ser considerado como sendo um sintoma dispéptico. Além disso, os achados da análise fatorial podem auxiliar no entendimento e futuros estudos dos diferentes mecanismos fisiopatológicos envolvidos na dispepsia funcional.

Dado os múltiplos fatores causais da dispepsia funcional, assim como pelas diversas facetas que influenciam na qualidade de vida, já era esperado que os sintomas dispépticos não tivessem uma relação estreita com a qualidade de vida. Contudo, a correlação negativa e estatisticamente significativa entre o PADYQ e o WHOQOL-BREF, corrobora a idéia de que o questionário seja um instrumento adequado, dada as diversas publicações mencionando inter-relação da dispepsia funcional e qualidade de vida, ainda que parcial. Achado interessante, ainda que dependendo de confirmação em outros trabalhos, é o achado de que o principal domínio que afeta a qualidade de vida dos pacientes com dispepsia funcional é o de Distensão/Saciedade.

Conclusões

O *Porto Alegre Dyspeptic Symptoms Questionnaire* (PADYQ) apresentou propriedades clinimétricas adequadas para uso em estudos clínicos. Mostrou ter adequada consistência interna e reprodutibilidade, tanto ao ser submetido a teste-reteste pelo mesmo entrevistador, quanto por entrevistadores diferentes. Também mostrou ter sensibilidade às mudanças acarretadas por a uma intervenção terapêutica e provou ser capaz de medir o constructo correto. Torna-se, assim, importante ferramenta de pesquisa validada para a investigação da dispepsia funcional. Não obstante, como foi construído a partir da terminologia utilizada internacionalmente e seus questionamentos não dependem do meio sócio-cultural em que está sendo utilizado, nós acreditamos que o PADYQ será facilmente traduzido e manterá suas propriedades clinimétricas quando utilizado em outros idiomas.

Referências bibliográficas

1. Talley N, Stanghellini V, Heading RC, Koch KL, Nalagelada J, Tytgat G. Functional gastroduodenal disorders. In *The Functional Gastrointestinal Disorders*, 2nd ed. Drossman DA, Corazziari E, Talley N, Thompson WG, Whitehead WE (eds). Lawrence, Degnon Associates, 2000 pp. 299–350
2. Holtmann G, Gschossmann J, Neufang-Huber J, Gerken G, Talley NJ: Differences in gastric mechanosensory function after repeated ramp distensions in non-consulters with dyspepsia and healthy controls. *Gut* 47(3):332–336, 2000
3. Mearin F, Cucala M, Azpiroz F, Malagelada JR: The origin of symptoms on the brain-gut axis in functional dyspepsia. *Gastroenterology* 101(4):999–1006, 1991
4. Fennerty MB: Is nonulcer dyspepsia related to *Helicobacter pylori* infection? *Semin Gastrointest Dis* 12(3):180–185, 2001
5. Talley NJ, Colin-Jones D, Koch K, Koch M, Nyren O, Stanghellini V: Functional dyspepsia: A classification with guidelines for diagnosis and management. *Gastroenterol Int* 4:145–160, 1991
6. Kirschner B, Guyatt G: A methodological framework for assessing health indices. *J Chronic Dis* 38(1):27–36, 1985
7. Sperber AD: Translation and validation of study instruments for cross-cultural research. *Gastroenterology* 126:s12.4–128, 2004
8. Streiner DL, Norman GR: *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, 1995
9. Cronbach L: Coefficient alpha and the internal structure of tests. *Psychometrika* 16(3):297–334, 1951
10. Streiner DL, Norman GR: Selecting the items. In *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, Publications, 1995.
11. Fachel J, Carney S: Avaliação psicométrica: A qualidade das medidas e o entendimento dos dados. In: Cunha J (ed). *Psicodiagnóstico V*. Porto Alegre, Artmed Editora, 2000
12. Shaw M, Talley NJ, Adlis S, Beebe T, Tomshine P, Healey M. Development of a digestive health status instrument: Tests of scaling assumptions, structure and reliability in a primary care population. *Aliment Pharmacol Ther* 12(11):1067–1078, 1998
13. O’Leary MP, Barry MJ, Fowler FJ Jr: Hard measures of subjective outcomes: Validating symptom indexes in urology. *J Urol* 148(5):1546–1548, 1992
14. Talley NJ, Verlinden M, Jones M: Quality of life in functional dyspepsia: Responsiveness of the Nepean Dyspepsia Index and development of a new 10-item short form. *Aliment Pharmacol Ther* 15(2):207–216, 2001
15. Streiner DL, Norman GR: *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, 1995

16. Liang MH: Evaluating measurement reponsiveness. *J Rheumatol* 22(6):1191–1192, 1995
17. Fleck M, Louzada S, Xavier M, Chachamovich E, Vieira RR, Santos L, Pinzon V: Aplicação da versão em português do instrumento de avaliação de qualidade de vida da Organização Mundial da Saúde (WHOQOL-100). *Rev Saude Publica* 33(2):198–205, 1999
18. Chassany O, Marquis P, Scherrer B, Read NW, Finger T, Bergmann JF, Fraitag B, Geneve J, Caulin C: Validation of a specific quality of life questionnaire for functional digestive disorders. *Gut* 44(4):527–533, 1999
19. Streiner DL, Norman GR: Validity: In *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, 1995
20. Fleck MP, Louzada S, Xavier M, Chachamovich E, Vieira G, Santos L, Pinzon V: [Application of the Portuguese version of the abbreviated instrument of quality life WHOQOL-bref]. *Rev Saude Publica* 34(2):178–183, 2000
21. Almeida TL: Anlise fatorial. <http://lula.dmat.furg.br/~taba/posanafat.htm>. Accessed 6 Aug 2003
22. van Zanten SJ, Talley NJ, Blum AL, Bolling-Sternevald E, Sundin M, Junghard O: Combined analysis of the ORCHID and OCA Y studies: Does eradication of *Helicobacter pylori* lead to sustained improvement in functional dyspepsia symptoms? *Gut* 50 (Suppl 4):iv26–iv30, 2002
23. van Zanten SJ: Assessment of outcome in dyspepsia: Has progress been made? *Gut* 50 (Suppl 4):iv23–iv25, 2002
24. Streiner DL: Sample size formulae for parameter estimation. *Percept Mot Skills* 78:275–284, 1994
25. van Zanten SJ, Tytgat KM, Pollak PT, Goldie J, Goodacre RL, Riddell RH, Hunt RH: Can severity of symptoms be used as an outcome measure in trials of nonulcer dyspepsia and *Helicobacter pylori* associated gastritis? *J Clin Epidemiol* 46(3):273–279, 1993
26. Buckley MJ, Scanlon C, McGurgan P, O’Morain CA: A validated dyspepsia symptom score. *Ital J Gastroenterol Hepatol* 29(6):495–500, 1997
27. Drossman DA: Chronic abdominal pain (with emphasis on functional abdominal pain syndrome). In Feldman, Sleisenger & Fordtran’s *Gastrointestinal and liver disease*, 7th ed. Feldman M, Scharschmidt BF, Sleisenger MH (eds). Philadelphia, W. B. Saunders, 2002
28. El Omar EM, Banerjee S, Wirz A, McColl K: The Glasgow Dyspepsia Severity Score—A tool for the global measurement of dyspepsia. *Eur J Gastroenterol Hepatol* 8(10):967–971, 1996

29. Junghard O, Lauritsen K, Talley NJ, Wiklund IK: Validation of seven graded diary cards for severity of dyspeptic symptoms in patients with non ulcer dyspepsia. *Eur J Surg Suppl* 583:106–111, 1998
30. Svedlund J, Sjdin I, Dotevall G: GSRS—A clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. *Dig Dis Sci* 33:129–134, 1988
31. Jaakkimainen RL, Boyle E, Tudiver F: Is *Helicobacter pylori* associated with nonulcer dyspepsia and will eradication improve symptoms? A meta-analysis. *BMJ* 319:1040–1044, 1999

ANEXO II - QUESTIONÁRIO DE SINTOMAS DISPÉPTICOS

Com relação aos últimos 30 dias

DOR

Qual a intensidade da dor abdominal (superior) na maioria dos dias neste período? ()

0. Ausente
1. Muito leve
2. Leve
3. Moderada
4. Forte
5. Muito forte

Qual a duração da dor na maioria dos dias neste período? ()

0. Não se aplica
1. Alguns minutos (menos que 30 minutos)
2. Menor que 2 horas
3. Maior que 2 horas

Com que frequência os Sr./Sra. apresentou dor abdominal nos últimos 30 dias? ()

0. Não se aplica
1. Raramente
2. 1 a 2 dias/semana
3. Quase diariamente
4. Diariamente

ESCORE TOTAL DOR _____ (máximo 12 pontos)

NÁUSEAS/VÔMITOS

Qual a intensidade das náuseas na maioria dos dias deste período? ()

0. Ausente
1. Muito leve
2. Leve
3. Moderada
4. Forte
5. Muito forte

Qual a duração aproximada da maioria dos episódios de náuseas? ()

0. Não se aplica
1. Alguns minutos (menos que 30 minutos)
2. Menor que 2 horas
3. Maior que 2 horas

Com que frequência o Sr./Sra. apresentou náuseas nos últimos 30 dias? ()

0. Não se aplica
1. Raramente
2. 1 a 2 dias/semana
3. Quase diariamente
4. Diariamente

Com que frequência o Sr./Sra. apresentou vômitos nos últimos 30 dias? ()

0. Não se aplica
1. Raramente
2. 1 a 2 dias/semana
3. Quase diariamente
4. Diariamente

ESCORE TOTAL NÁUSEAS/VÔMITOS: _____ (máximo 16 pontos)

DISTENSÃO/SACIEDADE

Qual a intensidade da sensação de distensão (“estufamento”/inchaço) nos últimos 30 dias? ()

0. Ausente
1. Muito leve
2. Leve
3. Moderada
4. Forte
5. Muito forte

Qual a duração destes episódios nestes períodos? ()

0. Não se aplica
1. Alguns minutos (menos que 30 minutos)
2. Menor que 2 horas
3. Maior que 2 horas

Com que frequência os Sr./Sra. apresentou esses episódios de distensão/inchaço no abdômen superior nos últimos 30 dias? ()

0. Não se aplica
1. Raramente
2. 1 a 2 dias/semana
3. Quase diariamente
4. Diariamente

Com que frequência o Sr./Sra. apresentou sensação de estar com o estômago cheio logo após começar a comer, nos últimos 30 dias? ()

0. Sem saciedade precoce
1. Raramente
2. 1 a 2 dias/semana
3. Quase diariamente
4. Diariamente

ESCORE TOTAL DISTENSÃO/SACIEDADE _____ (máximo 16 pontos)

PONTUAÇÃO TOTAL DOS SINTOMAS DISPÉPTICOS: _____ (máximo 44 pontos)

ANEXO III – VERSÃO PARA PUBLICAÇÃO EM INGLÊS

Development and Validation of a Brazilian Portuguese Questionnaire to Evaluate Functional
Dyspepsia

Guilherme Becker Sander, MD

Luiz Edmundo Mazzoleni, MD, ScD

Carlos Fernando Magalhães Francesconi, MD, ScD

André Castagna Wortmann, MD

Eduardo André Ott, MD

Alexandro Theil, MD

Vicenzo Da Cruz Piccoli

Ângela Cristiane Da Silva, MD

Leandro Oliveira, MD

Simone Beheregaray, MD

Simone Mاتیoti, MD

Gustavo Somm, MD

José Roberto Goldim, PhD

From the Graduate Program in Gastroenterology, Universidade Federal do Rio Grande do Sul,
Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil.

Address for reprint requests: Dr. Guilherme Becker Sander, Ramiro Barcelos, 1660 apto. 501,
CEP 90035–002, Porto Alegre, RS, Brazil; guilhermesander@hotmail.com.

Acknowledgements

We thank the following individuals for their contribution to this paper: Sérgio Gabriel Barros, MD, PhD, Júlio Pereira Lima, MD, PhD, Carisi Polanczyk, MD, ScD, Maria Helena Itaquí Lopes, MD, ScD, and Cléber Krueel, MD, PhD.

Abstract

Despite its high prevalence, nonulcer dyspepsia is still difficult to study, due to the lack of adequate tools to measure significant outcomes. The objective of this study was to develop and validate a symptom-focused, disease-specific questionnaire to evaluate patients with nonulcer dyspepsia. For that, the questionnaire was carefully written following widely accepted terminology, so as to facilitate translation and validation in other languages and cultures. The questionnaire was developed using Rome I terminology for symptoms, which were evaluated according to their intensity, duration, and frequency when applicable. Thirty-one patients with nonulcer dyspepsia, as well as 31 sex- and age-matched volunteers without digestive problems were used to assess the internal consistency, reproducibility, responsiveness, content validity, and discriminant validity of the questionnaire. Another 31 functional dyspeptic patients were enrolled for assessment of criterion validity. Cronbach's α coefficient was 0.82. The intraclass correlation coefficient for the scores obtained 7 days apart was 0.86. The mean score obtained after 3 months of treatment was 16.4, vs. 23.03 at baseline ($P = 0.001$). Two blinded gastroenterologists agreed that the questionnaire adequately evaluated nonulcer dyspepsia. The median symptoms score for controls was 0, vs. 22.5 for dyspeptic patients ($P = 0.001$). An inverse correlation was observed between quality of life and dyspeptic symptoms ($R = -0.28$, $P = 0.026$). The proposed questionnaire has high degrees of both reproducibility and responsiveness. As this questionnaire was based on Rome I International Consensus terminology, it is expected that it will be easy to translate and validate.

KEY WORDS: dyspepsia; questionnaires; questionnaire design; reproducibility of results.

Introduction

Dyspepsia is a very prevalent condition that affects about 25% of the world population, with an annual incidence of about 1%. It is defined as pain or discomfort located in the upper abdomen (1). The symptoms of dyspepsia include pain, postprandial fullness, early satiety, nausea, vomiting, and bloating. Most affected individuals do not present either structural or biochemical alterations that explain these symptoms and are, thus, classified as presenting nonulcer dyspepsia or functional dyspepsia (1). The pathophysiology of nonulcer dyspepsia is poorly understood; it may be associated with dysfunction in motility, secretion, and sensitivity of the digestive tract (2–4).

The diagnostic criteria for nonulcer dyspepsia were established by specialists in consensus statements known as the Rome criteria (1, 5). According to these criteria, nonulcer dyspepsia is diagnosed if the following symptoms are present for at least 12 weeks (not necessarily consecutive) during the past 12 months: (i) persistent or recurrent dyspepsia; (ii) no evidence of organic disease that is likely to explain the symptoms (including endoscopic evidence; and (iii) no evidence that relief of dyspeptic symptoms is associated exclusively with defecation or that dyspepsia is associated with changes in the frequency or stool form. Despite its great prevalence and the associated socioeconomic cost, nonulcer dyspepsia was not adequately studied until the end of the 20th century. This resulted in part from the dearth of available research tools. Being a disorder without defined structural or pathophysiological anomalies, the outcomes of studies focusing on the efficacy of nonulcer dyspepsia treatments are necessarily subjective. Thus, the development of methods capable of measuring symptoms and patients' perceptions in a reliable and reproducible manner becomes essential. This has been achieved through the development of standardized questionnaires.

When elaborating a questionnaire, it is necessary to select the most representative set of symptoms associated with the disorder to be investigated (6). To be practical, the

questionnaire must be planned to enable the largest amount of information to be obtained through the smallest possible number of questions. A questionnaire should also be sufficiently general to maintain its basic characteristics when used in different populations. Concerns about future translations should be kept in mind, as increasingly more multicentric studies are carried out. The challenge is to not create a questionnaire that is easy to translate but, rather, to assure comparability of results between different ethnic and cultural populations (7).

After it is devised, the questionnaire must be validated, that is, it must be tested to ensure that it is useful for clinical research. The properties of clinical measurement (known as clinimetric properties) that are considered essential for the validation of a questionnaire are internal consistency, reproducibility, responsiveness, and validity (8).

It is expected that the several items in a questionnaire will have at least a moderate level of relationship with one another and with the total score—so as to measure different aspects of the same condition, and not different characteristics of diverse entities. This attribute is known as internal consistency and can be estimated using Cronbach's coefficient (9)—a very high coefficient suggests redundancy, whereas a low coefficient suggests that the scale is evaluating more than one construct, the ideal values falling between 0.7 and 0.9 (10). Another tool used to accomplish this is factorial analysis. Factorial analysis has a fundamental role in empirically determining the dimensionality of items. A high degree of correlation between items suggests that all items measure the same concept. After this analysis, items found to have a low correlation with the total score may be excluded from the scale (11, 12).

Reproducibility aims at demonstrating that the data obtained are reliable. This feature can be measured by means of the “test-retest” method (13), according to which similar scores must be obtained in two distinct occasions without any treatment interventions between the two measurements. Testing and retesting are usually performed at an interval that is sufficiently short to ensure that there will be no changes to the underlying disease (14) while,

at the same time, sufficiently long to avoid memory bias (patient responding on the second measurement thinking about the answers given the first time around and not really reassessing the symptoms) (15).

Responsiveness expresses the sensitivity of the questionnaire to demonstrate changes if they actually occur (16). This property is tested by applying the instrument before and after a treatment intervention that has proven effects on the disorder under study.

Validity refers to the ability of an instrument to measure exactly what it is meant to measure (8). It can be assessed in different ways. Content analysis is not a statistical approach but, rather, a judgment by specialists in the field about representativeness and relevance of the items proposed in the scale (11). Another test is the extreme groups analysis—that is, by comparing a population with the disorder with a population without the disorder so as to assess the ability of the instrument to reveal the difference between the two populations. This kind of validity is also known as “discriminant validity” (17, 18). Another way to assess validity is by measuring “criterion validity.” In this case, an instrument of proven efficacy is applied in parallel with the instrument to be tested, and a good correlation must be obtained between the results of both questionnaires. If a scientifically established tool is not available, a related construct may be used, although in this case smaller correlation coefficients should be expected (18, 19).

The aim of our study was to develop and validate an easy-to-translate, symptom-focused (unidimensional) questionnaire to assess nonulcer dyspepsia. We chose a unidimensional questionnaire in this context for two reasons. First, quality-of-life questions are more susceptible to cultural adaptation errors in future translations. Second, the activity of nonulcer dyspepsia is strongly associated with quality of life. The fact that it is often difficult to determine the causal direction between these entities could undermine the use of quality of life as an outcome in nonulcer dyspepsia. Therefore, we developed an English questionnaire

based on the symptoms described as characteristic of nonulcer dyspepsia in the Rome I Consensus. To test the translatability and clinimetric properties of the questionnaire, it was translated to Brazilian Portuguese and validated in a group of Brazilian patients with nonulcer dyspepsia.

Methods

Development

The proposed Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) was developed based on the symptoms described as relevant in the Rome I consensus statement (5), resulting in an 11-item form (Table 1). All questionnaires were answered with the assistance of an interviewer. The patients were asked to average their symptoms during the preceding month.

Table 1 – Characteristic symptoms of nonulcer dyspepsia assessed by the Porto Alegre Dyspeptic Symptoms Questionnaire

Symptom	Score
Pain in upper abdomen	
Intensity	0-5
Duration	0-3
Frequency	0-4
Nausea	
Intensity	0-5
Duration	0-3
Frequency	0-4
Vomiting	
Frequency	0-4
Upper abdominal bloating	
Intensity	0-5
Duration	0-3
Frequency	0-4

Early satiety

Frequency

0-4

Total

0-44

Patients

Factorial analysis was performed taking into consideration all the baseline questionnaires of patients from a study about the effects of *H. pylori* eradication on nonulcer dyspepsia symptoms (n = 157). All patients were selected according to Rome I criteria.

To evaluate internal consistency, reproducibility and responsiveness, 31 patients with dyspeptic symptoms according to Rome I criteria were included in the study (Group A). All were submitted to upper digestive endoscopy at Hospital de Clinicas de Porto Alegre and had normal endoscopic results.

To determine discriminant validity, 31 healthy volunteers matched by sex and age to the group of patients described above and who reported having no digestive diseases were selected and interviewed after signing an informed consent form (control group).

To evaluate criterion validity, the control group and an additional group of 31 patients with nonulcer dyspepsia according to Rome I criteria (Group B) answered the Dyspeptic Symptoms Questionnaire and the World Health Organization Quality of Life (WHOQOL-BREF) (20) questionnaire.

The protocol was approved by the Research Ethics Committee at Hospital de Clínicas de Porto Alegre and all the patients and controls signed an informed consent form.

Evaluation of Clinimetric Properties

The clinimetric properties of the proposed PADYQ were evaluated in seven ways, as follows.

Factorial Analysis. Factorial analysis using the principal components method was used to aggregate items in subscales (domains). Components with *eigenvalues* larger than 1 were selected. After extraction of the components, orthogonal rotation of factors was performed using Kaiser's Varimax method. Items with a factorial load higher than 0.4 on the same component and which did not have values higher than 0.4 on other components were aggregated in domains (21).

Internal Consistency. Internal consistency was calculated using Cronbach's coefficient applied to the baseline questionnaires (TO) of Group A (9). The coefficient was expected to be higher than 0.7 but not higher than 0.9, which would mean redundancy (10, 12).

Reproducibility. The patients included in the study (Group A) answered the PADYQ during their first office visit (appointment TO). A new appointment was scheduled for 7 days later (appointment T1), and patients answered the questionnaire again. The patients were not exposed to any kind of intervention between T0 and T1. For 15 patients, the same interviewer completed the questionnaires at T0 and T1. For 16 patients, a different interviewer completed the two questionnaires, for evaluation of intra and inter-observer agreement. The correlation between scores was verified by calculating the intraclass correlation coefficient for both the total score and the different domains obtained by factorial analysis.

Responsiveness. Following the two initial evaluations (T0 and T1), the patients in Group A were randomly assigned to one of two groups: one group received 10-day courses of lansoprazole, 30 mg, clarithromycin, 500 mg, and amoxicillin, 1 g, all of them twice daily. The other group received lansoprazole, 30 mg twice a day, plus placebo tablets identical to the antibiotics in appearance and same posology. These groups were considered to be homogeneous, since there is no unequivocal evidence in the literature suggesting that one treatment has more efficacy than the other (22). A decrease of 13 to 73% in the scores was

expected, reflecting the response to placebo reported in previous studies (23). A new appointment (T2) was scheduled for 3 months after the beginning of treatment, and the PADYQ was again applied by one of the investigators. Responsiveness was calculated using the t test for paired samples, comparing the results of T1 and T2 for the total score and for the different domains assessed by the questionnaire.

Content Validity. Content validity was assessed by two experienced gastroenterologists blinded to the purpose of the questionnaire. They were asked to determine what construct the questionnaire was supposed to measure. Then they were asked to confirm that the questionnaire sampled the full range of symptoms of nonulcer dyspepsia.

Discriminant Validity. The baseline scores (T0) obtained for dyspeptic symptoms were compared with the scores obtained for the control group. This analysis was performed using the Wilcoxon test.

Criterion Validity. The PADYQ was applied in parallel to the WHOQOL-BREF (20) in 31 controls and in 31 patients with nonulcer dyspepsia (Group B). The results were correlated using Spearman' nonparametric coefficient of correlation.

Sample Size. The size of the sample required to test reproducibility was calculated expecting an intraclass correlation of 0.90 (considering an α error of 0.05 and aiming at confidence intervals of ± 0.10). According to Streiner (24), the number of subjects needed for this purpose was 31. Sample size for responsiveness was calculated to allow the use of paired t test comparison, considering an α error of 0.05, a power of 90%, an expected reduction of 20% in scores from an initial score of 25, a standard deviation of 8 at both visits, and a correlation of 0.5 among variables. The estimated sample size for responsiveness was 27. For factorial analysis, it is recommended to analyze at least five times more questionnaires than the number of items in the scale and suggested to analyze a minimum of 100 forms (21). For other clinimetric properties, sample size is difficult to estimate, so we used 31 patients as

calculated for reproducibility and responsiveness, in agreement with other studies of questionnaire validation that used similar samples (25, 26).

Results

Sample Characteristics

Of the 31 patients select, 24 were women. Sex-matching in the control group was perfect. The mean age of patients with dyspepsia was 43.1 years (Group A), and the mean age of controls was 43.0 (Pearson's correlation coefficient ≥ 0.99). Schooling was higher in controls than in patients with dyspepsia: 67% of the patients in the control group had finished high school, while 58% of the dyspeptic patients had less than 8 years of schooling.

Factorial Analysis

Among 11 items, factorial analysis identified 3 components with an *eigenvalue* higher than 1. The first component, with an *eigenvalue* of 3.077 and responsible for 24% of the total variance for the questionnaire, included the intensity, duration, and frequency of bloating, as well as the frequency of early satiety (factorial loads after Varimax rotation: 0.937, 0.876, 0.910, and 0.433, respectively). The second component, with an *eigenvalue* of 2.4 and responsible for 24.3% of the total variance, included the intensity, duration, and frequency of nausea and frequency of vomiting (factorial load after Varimax rotation: 0.912, 0.847, 0.873, and 0.577). The third component, with an *eigenvalue* of 2.224 and responsible for 21.1% of the total variance, included the intensity, duration, and frequency of upper abdominal pain (factorial load after Varimax rotation: 0.905, 0.841, and 0.867). None of the items had a factorial load higher than 0.2 on two components simultaneously (Table 2 and Figure 1). Thus, the first component was called the Bloating/Satiety domain; the second, the Nausea/Vomiting domain; and the third, the Pain domain. The sum of the variance for the three domains accounted for 70% of the variance for the questionnaire.

Table 2 – Factorial loads of the questionnaire items plotted after Varimax rotation.

	Component		
	1	2	3
Upper abdominal pain intensity	0,090	0,032	0,905
Upper abdominal pain duration	-0,036	0,070	0,841
Upper abdominal pain frequency	0,052	0,027	0,867
Nausea intensity	0,034	0,912	0,003
Nausea duration	-0,009	0,847	-0,025
Nausea frequency	0,072	0,873	0,021
Vomiting frequency	0,142	0,577	0,141
Upper abdominal bloating intensity	0,937	0,038	0,037
Upper abdominal bloating duration	0,876	0,004	-0,011
Upper abdominal bloating frequency	0,910	0,034	-0,074
Early satiety frequency	0,433	0,155	0,144

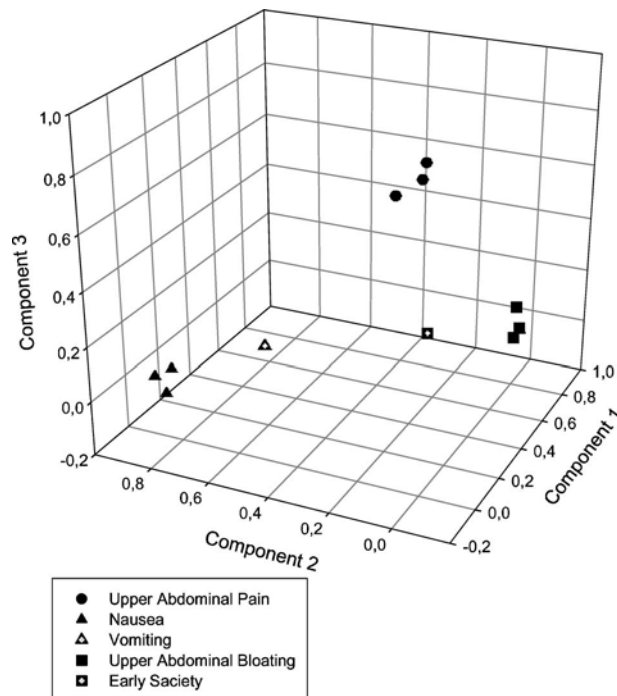


Fig 1. Factorial loads of the questionnaire items plotted after Varimax rotation.

Internal Consistency

Cronbach's α coefficient for the 11 items on the Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) answered by 31 patients with nonulcer dyspepsia (Group A) during appointment T1 was 0.82. Cronbach's α coefficient was 0.75 for the domain Bloating/Satiety, 0.90 for Nausea/Vomiting, and 0.83 for the Pain domain.

Reproducibility

The intraclass correlation coefficient for the total scores obtained at appointments T0 and T1 was 0.86 (Figure 2). The coefficient for interviews carried out by the same interviewer was 0.86; for interviews performed by different interviewers the coefficient was 0.87. The intraclass correlation coefficient was 0.87 for Bloating/Satiety, 0.88 for Nausea/Vomiting, and 0.68 for Pain. The mean of the total scores was 22.2 (SD, 1.56) for appointment T0 and 23.2 (SD, 1.42) for appointment T1 ($P \geq 0.2$).

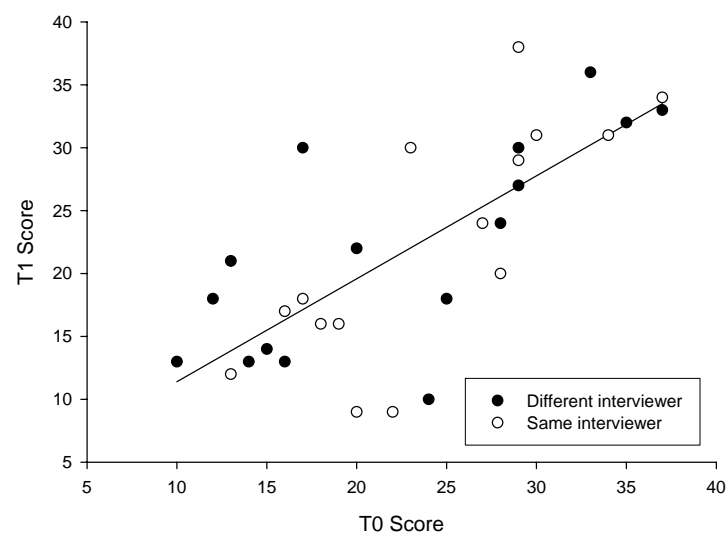


Figure 2. Correlation between scores obtained by applying the questionnaire at the T0 and T1 visits.

Responsiveness

As stated above, the mean T1 (pretreatment) score was 23.2. After drug treatment, at the 3-month follow-up visit, the mean score was 16.0 ($P \leq 0.001$). In patients receiving lansoprazole plus placebo ($n = 10$), the mean score was 18.8 (vs. a basal score of 25.9; $P = 0.038$). In patients receiving antibiotics ($n = 21$), the mean score after 3 months was 4.67 (vs. a basal score of 20.48; $P = 0.015$). The pretreatment score for Bloating/Satiety was 9.5, vs. 6.9 after 3 months ($P \leq 0.001$), for Nausea/Vomiting the mean score was 5.36 vs. 3.6 ($P = 0.082$), and for Pain the mean score was 7.4 vs. 5.5 ($P = 0.016$).

Content Validity

The two blinded gastroenterologists determined that the questionnaire evaluated nonulcer dyspepsia. They thought the items sampled the full range of symptoms of nonulcer dyspepsia and were relevant to this disease. The clarity of questions also was considered to be adequate.

Discriminant Validity

The mean Dyspeptic Symptoms score for controls was 0.9, with a median of 0 (minimum score = 0, maximum = 7) (Figure 3). The mean score for dyspeptic patients (Group A) was 22.2, with a median of 21.0 (minimum score = 9, maximum = 38). The difference between controls and dyspeptic patients was significant ($P \leq 0.001$). A Value of $P \leq 0.01$ was obtained for each of the comparisons between the results of the three domains in the group of controls and dyspeptic patients.

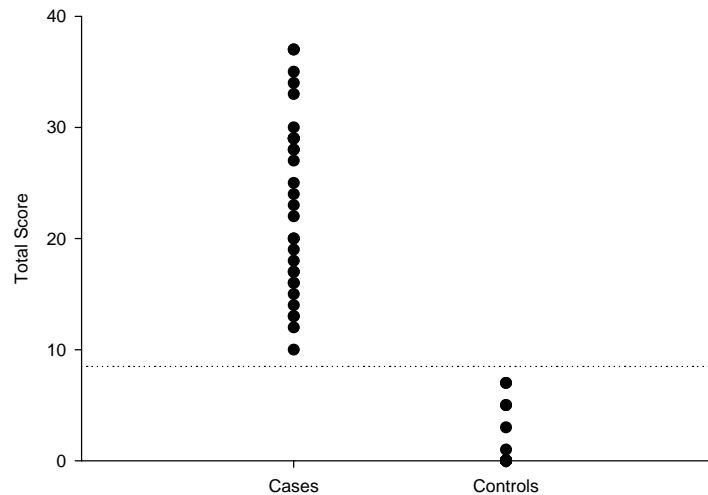


Figure 3 – Total score obtained with the questionnaire for patients with dyspepsia and controls.

Criterion Validity

There was a significant, although weak, correlation for the 31 patients with dyspepsia (Group B) and controls in terms of the scores obtained for the PADYQ and the WHOQOL-BREF ($R = -0.288$, $P = 0.023$). The correlation was negative: the higher the score obtained on the PADYQ, the lower the score on the WHOQOL-BREF. Concerning the domains of the PADYQ, the only domain presenting a significant correlation with the quality-of-life questionnaire was Bloating/Satiety ($R = -0.352$, $P = 0.005$).

Discussion

The questionnaire showed excellent clinimetric properties. Cronbach's alpha, which measures internal consistency (that is, the extent to which each item is related to other items and to the total score), was observed to be within the range considered as ideal (0.7–0.9) (15), both when the questionnaire's 11 items were analyzed together and when each domain was analyzed separately. The Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) was also shown to be reproducible when submitted to the test–retest procedure, both by the same interviewer and by different interviewers. Its responsiveness was shown to be adequate for clinical trials, since in groups as small as 10 patients the questionnaire was capable of detecting changes in symptoms resulting from the placebo effect. The PADYQ was also capable of adequately measuring the construct it aims to evaluate, since it was capable of markedly differentiating a group of patients without digestive diseases from a group with chronic symptoms in the upper abdomen. Content validity, evaluated through qualitative interviews, showed good item clarity and relevance.

Dyspepsia is a very prevalent problem, whose treatment still lacks efficacy (27). One of the problems that prevents the identification of efficacious treatments is the difficulty in measuring the benefits achieved when a new treatment modality is proposed. The Rome II consensus statement emphasizes the need to use assessment instruments that have been validated in clinical assays evaluating functional disorders of the digestive tract (1). The application of questionnaires that were created and validated in developed countries to different socioeconomic and cultural backgrounds in the developing world is questionable, especially when taking into account quality-of-life questions and measures of access to health-care facilities (28, 29).

Some existing questionnaires, in our opinion, present methodological flaws. The questionnaire developed by van Zanten et al. (25) addresses symptoms not currently

recognized as dyspeptic, such as heartburn, “sour taste,” bad breath, and eructation. A second important limitation of the validation process for that questionnaire was the performance of an upper digestive endoscopy between the first and the second applications of the questionnaire, during the period when reproducibility was being evaluated. The strong placebo effect of endoscopies, especially when the result is normal, is well known (22).

The questionnaire validated by Buckley et al. in 1997 (26) is also flawed in terms of the choice of symptoms reflecting dyspepsia and, also, includes heartburn and eructation, while other important dyspeptic symptoms such as nausea are not included. However, that questionnaire has the great advantage of detailing symptoms, focusing not only on the presence or absence of symptoms, but also on features such as duration, intensity, and frequency, which allows a more efficient quantification of nonulcer dyspepsia symptoms.

Another scale that has been used in large trials of nonulcer dyspepsia is the Gastrointestinal Symptom Rating Scale (GSRS). The GSRS uses a 7-point ordinal scale to measure a huge range of symptoms, not all considered part of nonulcer dyspepsia, such as abdominal pain, reflux, indigestion, diarrhea, and constipation. In fact, the scale was initially developed to be used with irritable bowel syndrome and peptic ulcer patients, not with nonulcer dyspepsia (30).

One of the most commonly used outcome measures in large studies of nonulcer dyspepsia is the Seven-Point Likert Scale (29). It is a self-completed instrument based on only one question: “Please state each day if you have experienced pain or discomfort in the stomach.” As such, the Likert Scale may be more liable to different interpretations by each patient, a shortcoming reflected in a lower than expected reliability coefficient (29, 31). This is especially true if the scale is applied to populations whose schooling level is different from that of the original population in which the questionnaire was validated. Also, this outcome measure evaluates the symptoms occurring within a 7-day interval, which may be too short to

assess a chronic disease whose symptoms vary in intensity (23). Finally, with the scoring system ranging only from 1 to 7, the discriminant power is very limited (31).

Other questionnaires (Glasgow score, for example) focus on the effect of dyspepsia on the life of patients, with little emphasis on dyspeptic symptoms (28). The Glasgow questionnaire was developed in Scotland, within an organized health system, and includes questions about the number of medical visits to the patient's home, number of medical consultations in the past 6 months, and use of nonprescription and prescription medication, which are likely to have a much smaller discriminant effect in other populations. Another problem is that responsiveness was determined in a study evaluating patients with peptic ulcer, and therefore, to our knowledge, the validation of the Glasgow score for functional dyspeptic patients has not been fully demonstrated (23).

When the PADYQ was created, the Rome II Consensus statement had not been released, so the questionnaire was based on Rome I criteria. In fact, the Rome II Consensus introduced few changes concerning the diagnostic criteria and range of symptoms considered to be associated with nonulcer dyspepsia. One notable exception is vomiting, present as one of the main dyspeptic symptoms in Rome I but not in Rome II. However, we found a strong correlation between vomiting and nausea in functional dyspeptic patients, corroborating the idea that vomiting should be considered a dyspeptic symptom. Additionally, factorial analysis findings could help in the understanding and further study of the different pathophysiologic mechanisms involved in nonulcer dyspepsia.

Taking into consideration the multiple causes of nonulcer dyspepsia, as well as the several aspects that influence quality of life, it was expected that dyspeptic symptoms would not be closely related to quality of life. However, the negative and statistically significant correlation between the PADYQ and the WHOQOL-BREF supports the notion that the questionnaire is an adequate instrument, since there have been several publications that report

an interrelation (although partial) between nonulcer dyspepsia and quality of life. Another interesting finding, which should be further investigated, is that the domain that most affects quality of life in nonulcer dyspepsia patients is that of Bloating/Satiety.

Conclusions

The Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) has adequate clinimetric features for use in scientific investigation. This questionnaire was shown to have adequate internal consistency and reproducibility when assessed by both the same and different investigators. It was also shown to be sensitive to changes following drug therapy and capable of measuring the correct construct. Therefore, it is an important research tool for the investigation of nonulcer dyspepsia. In addition, since it was based on internationally accepted terminology and on information that is not dependent on sociocultural background, we believe that the PADYQ is easy to translate and that it will maintain its clinimetric properties in other languages.

References

1. Talley N, Stanghellini V, Heading RC, Koch KL, Nalagelada J, Tytgat G. Functional gastroduodenal disorders. In *The Functional Gastrointestinal Disorders*, 2nd ed. Drossman DA, Corazziari E, Talley N, Thompson WG, Whitehead WE (eds). Lawrence, Degnon Associates, 2000 pp. 299–350
2. Holtmann G, Gschossmann J, Neufang-Huber J, Gerken G, Talley NJ: Differences in gastric mechanosensory function after repeated ramp distensions in non-consulters with dyspepsia and healthy controls. *Gut* 47(3):332–336, 2000
3. Mearin F, Cucala M, Azpiroz F, Malagelada JR: The origin of symptoms on the brain-gut axis in functional dyspepsia. *Gastroenterology* 101(4):999–1006, 1991
4. Fennerty MB: Is nonulcer dyspepsia related to *Helicobacter pylori* infection? *Semin Gastrointest Dis* 12(3):180–185, 2001
5. Talley NJ, Colin-Jones D, Koch K, Koch M, Nyren O, Stanghellini V: Functional dyspepsia: A classification with guidelines for diagnosis and management. *Gastroenterol Int* 4:145–160, 1991
6. Kirschner B, Guyatt G: A methodological framework for assessing health indices. *J Chronic Dis* 38(1):27–36, 1985
7. Sperber AD: Translation and validation of study instruments for cross-cultural research. *Gastroenterology* 126:s12.4–128, 2004
8. Streiner DL, Norman GR: *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, 1995
9. Cronbach L: Coefficient alpha and the internal structure of tests. *Psychometrika* 16(3):297–334, 1951
10. Streiner DL, Norman GR: Selecting the items. In *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, Publications, 1995.
11. Fachel J, Carney S: Avaliação psicométrica: A qualidade das medidas e o entendimento dos dados. In: Cunha J (ed). *Psicodiagnóstico V*. Porto Alegre, Artmed Editora, 2000
12. Shaw M, Talley NJ, Adlis S, Beebe T, Tomshine P, Healey M. Development of a digestive health status instrument: Tests of scaling assumptions, structure and reliability in a primary care population. *Aliment Pharmacol Ther* 12(11):1067–1078, 1998
13. O’Leary MP, Barry MJ, Fowler FJ Jr: Hard measures of subjective outcomes: Validating symptom indexes in urology. *J Urol* 148(5):1546–1548, 1992
14. Talley NJ, Verlinden M, Jones M: Quality of life in functional dyspepsia: Responsiveness of the Nepean Dyspepsia Index and development of a new 10-item short form. *Aliment Pharmacol Ther* 15(2):207–216, 2001

15. Streiner DL, Norman GR: Health Measurement Scales, 2nd ed. Oxford, Oxford Medical, 1995
16. Liang MH: Evaluating measurement responsiveness. *J Rheumatol* 22(6):1191–1192, 1995
17. Fleck M, Louzada S, Xavier M, Chachamovich E, Vieira RR, Santos L, Pinzon V: Aplicação da versão em português do instrumento de avaliação de qualidade de vida da Organização Mundial da Saúde (WHOQOL-100). *Rev Saude Publica* 33(2):198–205, 1999
18. Chassany O, Marquis P, Scherrer B, Read NW, Finger T, Bergmann JF, Fraitag B, Geneve J, Caulin C: Validation of a specific quality of life questionnaire for functional digestive disorders. *Gut* 44(4):527–533, 1999
19. Streiner DL, Norman GR: Validity: In Health Measurement Scales, 2nd ed. Oxford, Oxford Medical, 1995
20. Fleck MP, Louzada S, Xavier M, Chachamovich E, Vieira G, Santos L, Pinzon V: [Application of the Portuguese version of the abbreviated instrument of quality life WHOQOL-bref]. *Rev Saude Publica* 34(2):178–183, 2000
21. Almeida TL: Anlise fatorial. <http://lula.dmat.furg.br/~taba/posanafat.htm>. Accessed 6 Aug 2003
22. van Zanten SJ, Talley NJ, Blum AL, Bolling-Sternevald E, Sundin M, Junghard O: Combined analysis of the ORCHID and OCA Y studies: Does eradication of *Helicobacter pylori* lead to sustained improvement in functional dyspepsia symptoms? *Gut* 50 (Suppl 4):iv26–iv30, 2002
23. van Zanten SJ: Assessment of outcome in dyspepsia: Has progress been made? *Gut* 50 (Suppl 4):iv23–iv25, 2002
24. Streiner DL: Sample size formulae for parameter estimation. *Percept Mot Skills* 78:275–284, 1994
25. van Zanten SJ, Tytgat KM, Pollak PT, Goldie J, Goodacre RL, Riddell RH, Hunt RH: Can severity of symptoms be used as an outcome measure in trials of nonulcer dyspepsia and *Helicobacter pylori* associated gastritis? *J Clin Epidemiol* 46(3):273–279, 1993
26. Buckley MJ, Scanlon C, McGurgan P, O’Morain CA: A validated dyspepsia symptom score. *Ital J Gastroenterol Hepatol* 29(6):495–500, 1997
27. Drossman DA: Chronic abdominal pain (with emphasis on functional abdominal pain syndrome). In Feldman, Sleisenger & Fordtran’s *Gastrointestinal and liver disease*, 7th ed. Feldman M, Scharschmidt BF, Sleisenger MH (eds). Philadelphia, W. B. Saunders, 2002
28. El Omar EM, Banerjee S, Wirz A, McColl K: The Glasgow Dyspepsia Severity Score—A tool for the global measurement of dyspepsia. *Eur J Gastroenterol Hepatol* 8(10):967–971, 1996

29. Junghard O, Lauritsen K, Talley NJ, Wiklund IK: Validation of seven graded diary cards for severity of dyspeptic symptoms in patients with non ulcer dyspepsia. *Eur J Surg Suppl* 583:106–111, 1998
30. Svedlund J, Sjdin I, Dotevall G: GSRS—A clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. *Dig Dis Sci* 33:129–134, 1988
31. Jaakkimainen RL, Boyle E, Tudiver F: Is *Helicobacter pylori* associated with nonulcer dyspepsia and will eradication improve symptoms? A meta-analysis. *BMJ* 319:1040–1044, 1999

ANEXO IV – DYSPEPTIC SYMPTOMS QUESTIONNAIRE

Concerning the past 30 days

PAIN

How do you describe the intensity of upper abdominal pain on most of the days during this period? ()

0. Absent
1. Very mild
2. Mild
3. Moderate
4. Severe
5. Very severe

What has been the duration of pain on most of the days during this period? ()

0. Does not apply
1. A few minutes (less than 30 minutes)
2. Less than 2 hours
3. More than 2 hours

How often did you feel upper abdominal pain in the past 30 days? ()

0. Does not apply
1. Seldom
2. 1 to 2 days per week
3. Almost daily
4. Daily

TOTAL SCORE FOR UPPER ABDOMINAL PAIN: _____ (maximum 12 points)

NAUSEA

How intense is your nausea on most days during this period? ()

0. Absent
1. Very mild
2. Mild
3. moderate
4. Severe
5. Very severe

What has been the duration of most nausea episodes? ()

0. Does not apply
1. A few minutes (less than 30 minutes)
2. Less than 2 hours
3. More than 2 hours

How often did you feel nausea in the past 30 days? ()

0. Does not apply
1. Seldom
2. 1 to 2 days per week
3. Almost daily
4. Daily

VOMITING

How often did you vomit in the past 30 days? ()

- 0. Does not apply
- 1. Seldom
- 2. 1 to 2 days per week
- 3. Almost daily
- 4. Daily

TOTAL SCORE FOR NAUSEA/VOMITS: _____ (maximum 16 points)

UPPER ABDOMINAL BLOATING

How intense has been your feeling of upper abdominal bloating (distension sensation/fullness) in the past 30 days? ()

- 0. Absent
- 1. Very mild
- 2. Mild
- 3. Moderate
- 4. Severe
- 5. Very severe

How long have these episodes lasted during this period? ()

- 0. Does not apply
- 1. A few minutes (less than 30 minutes)
- 2. Less than 2 hours
- 3. More than 2 hours

How often have you experienced upper abdominal bloating/fullness in the past 30 days? ()

- 0. Does not apply
- 1. Seldom
- 2. 1 to 2 days per week
- 3. Almost daily
- 4. Daily

EARLY SATIETY

In the past 30 days, how often have you felt that your stomach is full right after you start eating? ()

- 0. Does not apply
- 1. Seldom
- 2. 1 to 2 days per week
- 3. Almost daily
- 4. Daily

TOTAL SCORE FOR UPPER ABDOMINAL BLOATING/EARLY SATIETY: _____ (maximum 16 points)

TOTAL SCORE FOR DYSPEPTIC SYMPTOMS: _____ (maximum 44 points)

ANEXO V - GRÁFICOS ADICIONAIS

Figura 3 - "Scree Plot" - Valores de raízes latentes dos diferentes componentes

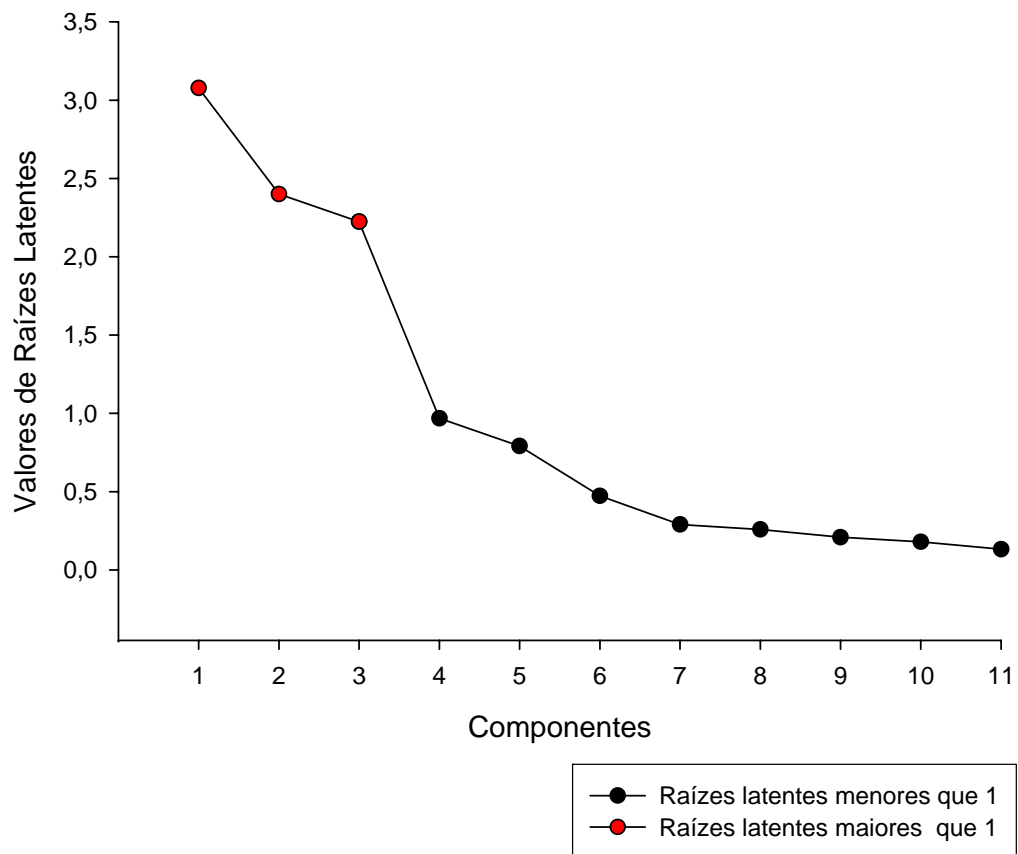


Figura 4 - Cargas fatoriais dos itens do questionário plotados após rotação varimax

