



The
University
Of
Sheffield.

School of
Health
And
Related
Research

HEALTH ECONOMICS & DECISION SCIENCE

Discussion Paper Series

Title: Assessment of the psychometric properties and refinement of the Health and Self-Management in Diabetes Questionnaire (HASMID)

Authors: Jill Carlton, Donna Rowen, Jackie Elliott

Corresponding author: Jill Carlton, School of Health and Related Research (SchARR) Regent Court, 30 Regent Street, Sheffield S1 4DA

**Tel: (+44) (0)114 222 5454 Fax: (+44) (0)114 222 0785
No. 18.08**

Disclaimer:

This series is intended to promote discussion and to provide information about work in progress. The views expressed in this series are those of the authors. Comments are welcome, and should be sent to the corresponding author.

This paper is also hosted on the White Rose Repository: <http://eprints.whiterose.ac.uk/>

**Assessment of the psychometric properties and refinement of the Health and Self-Management in
Diabetes Questionnaire (HASMID)**

Jill Carlton (PhD)¹, Donna Rowen (PhD)¹, Jackie Elliott (PhD)^{2,3}

¹ Health Economics and Decision Science, School of Health and Related Research, University of Sheffield,
Sheffield, UK

² Academic Unit of Diabetes, Endocrinology and Metabolism, Department of Oncology & Metabolism,
University of Sheffield, Sheffield, UK

³ Sheffield Teaching Hospitals NHS Trust, Diabetes and Endocrine Centre, Northern General Hospital, Sheffield,
UK

Corresponding author:

Jill Carlton

School of Health and Related Research (SchARR)

University of Sheffield

Regent Court, 30 Regent Street

Sheffield, UK

S1 4DA

j.carlton@sheffield.ac.uk

*

ABSTRACT

Background

The Health And Self-Management In Diabetes (HASMID^{v1}) questionnaire consists of 8 attributes, 4 about quality of life, and 4 about self-management. The overall aim of this study was to rigorously examine the psychometric properties of the HASMID^{v1} questionnaire.

Methods

The study comprised two phases. Phase 1 identified items of the HASMID^{v1} questionnaire that potentially required rewording through consultation with a patient involvement panel and two focus groups with people with diabetes. Phase 2 involved a cross-sectional longitudinal survey where HASMID, EQ-5D-5L, health, treatment and sociodemographic questions were administered using both paper and online versions to people with diabetes. Participants were asked to complete the survey again approximately 3 months later. Psychometric analyses were undertaken to examine floor and ceiling effects, item distributions, known group differences, change between time points, and responsiveness defining change using self-reported general health and self-management. Rasch analysis was undertaken to assess differential item functioning and disordered thresholds.

Results

Phase 1 derived five alternative items: Irritable, Affects Mealtimes, Daily Routine, Social Activities and Problem. Phase 2 achieved 2835 responses in time point 1 (n=1944 online, n=891 paper version) and 1243 in time point 2 (n=533 online, n=710 paper version). Overall the HASMID items performed well, though two alternative items (Irritable and Social Activities) provided additional information not fully captured by the original HASMID items.

Conclusion

Psychometric evaluation and Rasch analysis were used in conjunction with expert opinion to determine the final questionnaire. An additional two items (Irritable and Social Activities) have now been added to the original HASMID instrument. The application of psychometric analyses or Rasch analysis alone to inform item selection would have resulted in different items being selected for the final instrument. The benefit of a combined approach has produced an instrument which has a broader evaluation of self-management. The final validated HASMID-10 is short self-report PRO that can be used to evaluate the impact of self-management for people living with diabetes. HASMID-10 can be scored using total summative scores, with utility and monetary values also available for use in cost-utility and cost-benefit analyses.

Keywords

Diabetes mellitus

Self-management

Patient reported outcome (PRO)

Psychometric

Quality of life (QoL)

Questionnaire

Assessment of the psychometric properties and refinement of the Health and Self-Management in Diabetes Questionnaire (HASMID)

BACKGROUND

Diabetes is a complex condition in which the long-term consequences of the disease are largely determined by the ability of the individual to self-manage dietary habits, physical activity and various medications. In terms of glucose control, people with diabetes (PwD) need to avoid chronic hyperglycaemia (high glucose levels in the bloodstream) in order to evade microvascular diabetes-related complications which include eye damage (retinopathy), kidney damage (nephropathy) and neuropathy (nerve damage), which if severe can lead to blindness, dialysis and leg amputations, respectively. Macrovascular complications such as heart attacks and strokes are also increased as a result of chronic hyperglycaemia. Some medications can cause hypoglycaemia (low blood glucose levels in the bloodstream), which if mild may cause for example sweating, slurred speech and tingling, but if more severe can cause confusion, loss of consciousness, seizures and occasionally may be fatal. Many people with type 2 diabetes (T2DM) control their glucose through adapting their diet or taking tablets that do not cause hypoglycaemia. For some, with longer duration of T2DM, medications may be required which if not given at just the right dose may cause hypoglycaemia (including insulin). For people with type 1 diabetes (T1DM), insulin treatment is the only option, and doses need to be adjusted according to the amount of carbohydrate in each meal/drink (as this is the primary source of glucose), the current glucose level, the intensity of physical activity (prior and post the injection) and whether or not there has been any recent alcohol consumption. As there are several factors to take into consideration multiple times each day it is not unsurprising that avoiding hyper- and hypoglycaemia is more challenging in T1DM than in T2DM. In all forms of the disease (T1DM accounts for approximately 5%, T2DM approximately 90%, other approximately 5%), other aspects of diabetes care may include medication for raised blood pressure and/or raised cholesterol, as well as therapy for the diabetes-related complications. Hence, the ability of the patient to self-manage their condition is of

paramount importance to help them live as healthy a life as possible. Some PwD live a normal life-span with minimal complications of diabetes, whereas others suffer devastating complications and die early [1]. Therefore, educational interventions which aim to improve self-management skills such as DESMOND for T2DM, and DAFNE [2] for T1DM, are seen as the cornerstone of good diabetes care [3].

In a previous study we confirmed that self-management plays a significant role in the life of PwD. During the project we developed a patient reported outcome (PRO) measure that was able to measure the quality of life impact of self-managing the condition capturing both health and treatment experience. The Health And Self-Management In Diabetes (HASMID^{v1}) questionnaire consists of 8 attributes, 4 about quality of life, and 4 about self-management [4]. However, psychometric analyses around the performance of the questionnaire have not been previously undertaken. The validation of all PROs is important to demonstrate that they are able to measure what they are intended to measure. Validity, reliability and responsiveness of PROs should be assessed and reported so that users can be assured that the instrument is appropriate to use. The overall aim of this study was to rigorously examine the psychometric properties of the HASMID^{v1} questionnaire. The study had two phases: Phase 1 concerned the identification of any items of the HASMID^{v1} questionnaire that may require rewording and deriving alternative and any additional items; Phase 2 concerned the assessment of the psychometric properties of the HASMID^{v1} questionnaire in a large observational survey of people with diabetes using two methods of administration (online and postal).

METHODS

HASMID^{v1} questionnaire

The questionnaire consists of 8 attributes with three response options (never, sometimes, usually and always). Response options are scored from 0 to 3 with a higher score indicating little or no impact upon quality of life. The overall questionnaire is scored from 0 to 24, with a high score indicating good health related quality of life, and a low score indicating poor health related quality of life [4]. The

measure has been valued using discrete choice experiments that enable the measure to be used to generate QALYs for use in cost-utility analyses or willingness-to-pay values to use in cost-benefit analyses [5,6].

[Insert Figure 1 here]

Phase 1

The aim of Phase 1 was to critically evaluate the wording of the HASMID^{v1} questionnaire, and to consider whether any of the existing items could be rephrased. This was undertaken through presentation of the HASMID^{v1} questionnaire to a Patient and Public Involvement panel, who advised on aspects of wording on the questionnaire, and then two focus groups were conducted with PwD. The inclusion criteria were that the participant: had a clinical diagnosis of diabetes, was aged 18 years or greater; fluent in English; and able to provide informed consent. Potential participants were identified through an existing University research database. At the focus groups participants were asked to comment upon a number of aspects of the HASMID^{v1} questionnaire. This included the wording of the instructions for completing the questionnaire; consideration of the wording of existing items; and suggestions (if any) for refining items. Participants were also asked if there were any other aspects of living with diabetes that were not covered by the HASMID^{v1} questionnaire. Feedback was also requested on the layout of the questionnaire itself, including size and style of font. Based on the comments from the focus group additional items were formed to test in Phase 2. Alternatively phrased items were generated for mood (one alternative item), social limitations (three alternative items) and hassle (one alternative item) (see Figure 1). No new items were suggested.

Phase 2

The aim of Phase 2 was to evaluate the psychometric properties of the HASMID^{v1} questionnaire, including sensitivity and responsiveness, and to determine whether the alternatively phrased items

identified in Phase 1 performed better than the original HASMID^{v1} items, suggesting whether any amendments are necessary to HASMID^{v1}.

Recruitment

This project sought to formally test the developed HASMID^{v1} questionnaire (with the alternatively phrased items generated in Phase 1). We conducted a longitudinal survey with PwD (including both Type 1 and Type 2). Both online and paper versions of the questionnaire were tested. Potential participants were recruited via four main cohorts:

- DAFNE Online (see <http://www.dafneonline.co.uk/>), a panel of over 1,500 DAFNE graduates; a website designed specifically for people with T1DM who have undertaken a DAFNE structured education, but also accessible to anyone wishing to find out more about T1DM. Recruited via online link providing direct access to the online survey
- Diabetes UK (see <http://diabetes.org.uk/>); the main charity for all patients with diabetes in the UK. Members received an electronic link in their online newsletter and a printed link on their printed newsletter to the online survey
- A panel of over 2,300 patients at Sheffield Teaching Hospitals NHS Trust who consented to be contacted for research studies in diabetes. Potential participants were randomly allocated to participate in either the online or paper-version of the survey. Postal participants were sent information sheets and questionnaire via post and returned this in a pre-paid envelope. Online participants were sent an Invitation to Participate letter providing information on how to access the online survey.
- Social media through Twitter and a University of Sheffield mailing list, with a link to the online survey.

All consenting participants were given the option to enter a prize draw, with one £50 voucher randomly selected for each 50 participants.

Data collection

All participants were asked to complete the HASMID^{v1} questionnaire, EQ-5D-5L, and sociodemographic and health questions. Data was collected at more than one time point in order to allow for testing of responsiveness. Both online and postal respondents were asked if they would be willing to complete the survey again in approximately 3 months' time, for another chance of "winning" a £50 shopping voucher. Those who completed the postal survey were sent an additional pack by Sheffield Teaching Hospitals NHS Trust. Those who completed the online survey were sent a reminder email.

Analysis

Standard descriptive analysis was undertaken on the sample, with sub analysis on mode of administration of the survey. The psychometric properties of the HASMID questionnaire (and alternatively phrased items) were explored. Consideration of acceptability, reliability and validity were directed in the analysis.

Floor and ceiling effects

Floor and ceiling effects describe the amount of responses given at either end of the scale. A high percentage of floor or ceiling could be suggestive that the content validity of the instrument is limited. There is no universal definition of acceptable values for floor and/or ceiling effects. A 15% threshold was adopted for this study in line with other studies [7-11].

Responsiveness

The proportion of respondents reporting a change across the two time points was assessed for each item according to the level of the change (-3, -2, -1, 0, 1, 2, 3). Responsiveness was assessed using change across the two time points with the self-report general health and general self-management questions used to indicate where there was a change across the two time points. For this study, a mean change of >0.1 was considered as significant responsiveness.

Test-retest reliability

Test-retest reliability was determined by a mean change of <0.1 between time point 1 and time point 2, for respondents reporting no change in their general health and no change in their self-management.

Validity

Validity was assessed by examining known-group differences for each item across T1DM and T2DM respondents. The difference in the means between groups, effect size (using Cohen's d) and T-tests were undertaken to assess item performance. The standard effect size value can be described to fall within the ranges of small 0.2 to 0.5, medium 0.5 to 0.8 and large >0.8 . A value of <0.2 is considered nonsignificant [12]. It was hypothesised that respondents with T1DM would report lower QoL than respondents with T2DM, as previously reported with other measures [13].

Rasch analysis

Rasch analysis is a logit modelling technique that can be used to inform the selection of items in a patient reported outcome measure [14] and can be used to select best performing items (for example 15-18). Rasch analysis was used to assess differential item functioning (DIF). DIF is a form of bias across groups of respondents occurring when different groups within the same sample (despite equal severity of the underlying characteristic of health/self-management) respond in a different manner to an individual item. Rasch analysis was applied to identify whether DIF was present by gender, age,

and type of diabetes mellitus (DM). DIF by type of DM may not be a reason for removing an item, but is indicative that the item performs differently across T1DM and T2DM. For this study Rasch techniques were applied to a subset of randomly selected participants (n=500) for each mode of administration and for both time points. The analysis was conducted separately for each of the following scenarios; comparing Temper and Irritable; comparing Hassle and Problem; comparing Tied Mealtimes, Affects Mealtimes, Daily Routine and Social Activities. For completeness the remaining original HASMID^{v1} items were also considered and included in one analysis (Hypo, Tired, Control, Stressful and Support).

Item selection

The results of the psychometric evaluation and Rasch analysis were used to inform final item selection for the HASMID questionnaire. Alternatively phrased items were directly compared to the original item, such that: Temper versus Irritable; Hassle versus Problem; and Tied Mealtimes versus Affects mealtimes, Daily Routine, Social Activities. Item selection was informed by consideration of floor and ceiling effects; responsiveness and test-retest reliability; DIF; and clinical opinion.

RESULTS

Phase 2

The samples

A total of 2835 participants completed the survey. A larger proportion of participants completed the survey online (69%). Table 1 details the differences between the two samples. There was a higher percentage of male respondents in the postal survey (at both time points), whereas there was a higher percentage of female respondents in the online survey (at both time points). The mean age of respondents was lower for online respondents (at both time points). There were a higher percentage of T1DM respondents in the online survey compared to the postal survey (for both time points). The samples are very different for both time points and mode of administration (i.e. postal or online

survey). The online sample are younger, more likely to be female, more likely to be in employment, and more likely to rent (compared to postal respondents). The proportion of people with T1DM and T2DM differed by sample, where the proportion of T1DM is much lower in the postal samples. It was therefore necessary to undertake separate psychometric analyses for the four samples: online time point 1, postal time point 1, online time point 2 and postal time point 2. Where relevant, the analyses were also separated by people with T1DM and T2DM across each of the 4 samples. The HASMID^{v1} scores were derived from the original eight HASMID^{v1} items. The HASMID^{v1} questionnaire appears to suggest that there is a difference in scores from T1DM and T2DM respondents. T1DM reported lower QoL compared to T2DM.

[INSERT TABLE 1 HERE]

Responsiveness

Responsiveness was assessed using change across the two time points using the self-report general health and general self-management questions to indicate where there was a change across the two time points, summarised in Table 2. A tick indicates a mean change ≥ 0.1 in the same direction as the change in general health/self-management for both an improvement and deterioration in general health/self-management. Some items demonstrated a mean change in the right direction for an improvement in general health/self-management. Some items demonstrated a mean change in the right direction for a deterioration of general health/self-management. A cross indicates this criteria was not met. As the survey was not attached to a clinical intervention, there were no expectations for any individual to experience a change in the follow up period, and any change is subjectively reported by the respondent. (Detailed results are shown in Tables 1 and 2 in Additional Material).

[INSERT TABLE 2 HERE]

Where there is a change in general health, the health items perform as expected with typically a reduction in mean item response when general health has worsened and an increase in mean item response when general health has improved. Equally where there is a change in general self-management, the self-management items perform as expected with typically a reduction in mean item response when general self-management has worsened and an increase in mean item response when general self-management has improved. No single item demonstrated “acceptable” levels of responsiveness for both self-report general health and self-report self-management. When considering only general health Tired fulfilled the criteria for both T1DM and T2DM groups. When considering only self-management, Control fulfilled the criteria for both T1DM and T2DM groups.

Test-retest reliability

There is little change for most of the HASMID when there is no change in either general health or general self-management between time point 1 and time point 2, which suggests good test-retest reliability for these items. Irritable, Hypo, Tired, Tied Meal Times, Affects Meal Times, Control, Daily Routine, Hassle and Problem all fulfilled the criteria for both T1DM and T2DM groups for general health/self-management. (Detailed results are shown in Tables 1 and 2 in Additional Material).

Validity

Validity of the HASMID items (original and alternatively phrased items) was assessed by examining known-group differences for each item across T1DM and T2DM respondents. The summarised results of floor and ceiling effects are shown in Table 5. (Detailed results are shown in Tables 3 to 6 in Additional Material). Each of the samples shows different results. All items demonstrated floor and/or ceiling effects for at least one time point and/or mode of administration. No item demonstrated floor and/or ceiling effects for every time point across every sample.

[INSERT TABLE 3 HERE]

Each of the samples shows different results for known-group differences. For the online sample at time point 1 all items except Daily Routine demonstrate statistically significant differences between T1DM and T2DM respondents. However, the effect sizes of some of the items are non-significant (Temper, Irritable, Tired, Tied Mealtimes, and Control). For the postal sample at time point 1 all items except Control and Support demonstrate statistically significant differences between T1DM and T2DM respondents. Non-significant effect sizes were also seen for these items (Control and Support). For the online sample at time point 1 the following items did not show statistically significant differences between T1DM and T2DM respondents: Temper, Tired, Tied Mealtimes, and Support. Non-significant effect sizes were seen for Temper, Tired, Tied Mealtimes, Control, and Support. For the postal sample at time point 2 the following items did not show statistically significant differences between T1DM and T2DM respondents: Temper, Irritable, Tired, Tied Mealtimes, Control, Social Activities, and Support. Non-significant effect sizes were seen for the items Temper, Irritable, Tired, Tied Mealtimes, Control, Social Activities, and Support. Hypo, Affects Mealtimes, Hassle, Problem and Stressful demonstrated statistically significant differences between T1DM and T2DM respondents at both times points, and for both modes of administration.

Summary of psychometric properties

No single item demonstrated “acceptable” levels of responsiveness and test-retest reliability for both self-report general health and self-report self-management. When considering responsiveness as assessed by a change in general health, Tired, Daily Routine, Stressful and Support fulfilled the criteria. When considering only self-management, Control fulfilled the criteria. All items demonstrated floor and/or ceiling effects for at least one time point and/or mode of administration. No item demonstrated floor and/or ceiling effects for both time points across each sample. Hypo, Affects Mealtimes, Hassle, Problem and Stressful demonstrated statistically significant differences between T1DM and T2DM respondents at both times points, and for both modes of administration.

Summary of Rasch analysis

Full results of the DIF analyses can be found in Tables 7-18 in Additional Material. A summary of the Rasch analyses are shown in Table 4.

[INSERT TABLE 4 HERE]

Comparing Temper and Irritable

The results of the analysis across the different samples showed different results. Temper demonstrated DIF for age for the postal and online samples at time point 2, and DIF for gender for the postal sample at time point 1. Irritable showed DIF for age for the postal sample at time point 1 and online sample at time point 2, as well as DIF for gender for the postal sample at time point 1. Neither item showed disordered thresholds for any of the samples.

Comparing Hassle and Problem

The results of the analysis across the different samples showed different results. Hassle demonstrated DIF for gender, age and DM type for the online sample at time point 2, as well as DIF for gender for the postal sample at time point 2. Problem demonstrated DIF for gender and DM type for the online sample at time point 2. Neither item showed disordered thresholds for any of the samples.

Comparing Tied Mealtimes, Affects Mealtimes, Daily Routine and Social Activities

The results of the analysis across the different samples showed different results. Tied Mealtimes showed DIF for age across all samples (mode of administration and time points). It also showed DIF for gender and DM type for the online sample at time point 2. Affects Mealtimes showed DIF for gender across three of the samples (except online sample at time point 1). Daily Routine showed DIF for DM type across all of the samples, and DIF for age for three of the samples (except online sample

at time point 1). It also showed disordered thresholds for the postal sample at time point 2, and online sample at time point 1. Social activities demonstrated disordered thresholds for the postal sample at time point 2.

Examining Hypo, Tired, Control, Stressful and Support

The results of the analysis across the different samples showed different results. Hypo showed disordered thresholds and DIF for DM type across all of the samples. It also showed demonstrated DIF for gender for the postal sample at time point 2. Tired demonstrated DIF for age for the postal sample at time point 1. It also showed DIF for DM type for the online samples at both time point 1 and time point 2. Control demonstrated DIF for gender for the postal samples at time point 1 and time point 2, and disordered thresholds for the postal sample at time point 1. Stressful demonstrated DIF for age for the postal sample at time point 1 and the online sample at time point 1. It also demonstrated DIF for DM type for the online sample at time point 1. Support demonstrated DIF for DM type for the online sample at time point 1, and DIF for age for the online sample at time point 2.

Final item selection

Item selection was informed by the analyses reported above. Firstly, consideration was made on the alternative items. Finally the remaining original HASMID items were discussed to determine whether there was sufficient evidence to remove any items from the questionnaire.

Temper versus irritable

Conceptually both items can be considered as “mood” items, with temper tapping into the more severe end of the spectrum. Having items that can measure at the extremes can be useful. The psychometric results showed both items were moderately correlated. Differences were noted in the floor and ceiling effects (particularly across different time points and groups). However it is important to remember that the groups are different in their background characteristics. The responsiveness

analysis demonstrated that irritable does not work in the way we would expect, and temper is able to demonstrate a better responsiveness rate. Both items also demonstrated DIF for gender and age. Clinical opinion was considered, and whilst it was acknowledged that respondents are more likely to be aware of if they have experienced temper (rather than irritability), temper itself could be considered as a personality trait. Similarly, there are possible social connotations of admitting to having experienced temper. Irritable is a milder item, that is unlikely to be a personality trait and there are no social connotations of feeling/being irritable. A decision was made to retain Temper as an item, and to include Irritable as an additional item.

Hassle versus problem

Both items appeared to perform similarly psychometrically, with similar levels of responsiveness, floor/ceiling effects and DIF. Given that there is no overwhelming evidence to support the notion to change the item from hassle to problem, the original wording was kept.

Tied Mealtimes versus Affects Mealtimes, Daily Routine, Social Activities

All four items appeared to have similar psychometric properties in terms of floor and ceiling effects and responsiveness. There were issues with DIF for each item, particularly with respect to age. Out of the four items, social activities item was preferred as it demonstrated no DIF for age. From a clinical perspective there were concerns around the rigidity of concepts such as mealtimes and routine. It was felt that whilst these may be an issue to some, these could be generational concepts where older people may be more likely to follow a routine and have stricter mealtimes (as shown by the DIF analysis). In the longer term these items could be redundant – people with DM now may not have a mealtime/routine in the way that someone of an older generation has. A decision was made to include Social Activities as an additional item.

Hypo, Tired, Control, Stress, and Support

The Hypo item is more relevant to people with T1DM as those with T2DM. Therefore, not unsurprisingly the floor and ceiling effects in psychometric analyses were better for those with T1DM as opposed to T2DM. This is reflected in the Rasch analysis as it demonstrated DIF for DM type. It's inclusion in the final questionnaire is driven by the importance of the concept itself. The fear of going hypo will be more prevalent in those with T1DM, as virtually all people living with T1DM will have experienced hypoglycaemia, and have to regularly self-adjust their medication to minimise the chances of hypoglycaemia, whilst also avoiding hyperglycaemia, in order to control their HbA1c levels and avoid long-term complications of diabetes. The same cannot be said for the majority of people living with T2DM. As whilst some people living with T2DM will have to adjust their medication, under direction of their GP / physician, to control their HbA1c levels; the majority will either follow a controlled diet alone, or a controlled diet with a tablet medication that does not put them at risk of hypoglycaemia. Only those people with more complex T2DM will need medication that can cause hypoglycaemia (if given inappropriately). Thus, the level of self-management engagement required for glucose control, to avoid hypo- and hyper-glycaemia, in people with T2DM is lower than T1DM.

Tired demonstrated acceptable responsiveness and test-retest reliability for self-reported general health. It fulfilled the criteria for test-retest reliability for self-reported self-management. There were issues with floor effects for three of the 4 samples (online and time point). It also demonstrated some DIF for DM Type across two samples (online). Control fulfilled the criteria for test-retest reliability for both general health and self-management. It met the criteria for responsiveness for self-management; however, it did not perform as well for general health. There were issues with floor effects for all of the 4 samples. It demonstrated DIF for gender across two of the samples (postal).

Stressful fulfilled the criteria for test-retest reliability for general health; however it did not perform as well for responsiveness for either general health or self-management. There were issues with ceiling effects for some samples (time point and mode of administration), and it demonstrated DIF for age for two samples (time point 1). Support fulfilled the criteria for test-retest reliability for general health; however it did not perform as well for responsiveness for either general health. There were

issues with ceiling effects for some samples, and it demonstrated DIF for some samples. A decision was made to retain all items given insufficient evidence to suggest their removal. All concepts were felt to be clinically important.

HASMID-10

The final measure, HASMID-10 (Figure 2), consists of ten items that cover Temper, Irritable, Hypo, Tired, Tied mealtimes, Social Activities, Control, Hassle, Stress, and Support (the original HASMID^{v1} items plus Irritable and Social Activities). The response options for the HASMID-10 are that of the original PRO (never, sometimes, usually, and always). The questionnaire is scored summatively, with response levels being scored as never=3, sometimes=2, usually=1, always=0. Scores can range from 0-30, with a higher score indicating better quality of life. Utility scores and willingness to pay values can be generated for the measure using the original eight item HASMID^{v1} questionnaire (Rowen et 2017; 2018).

DISCUSSION

In the development of any PRO it is important to fully evaluate its performance prior to mainstream usage. The present study has allowed us to further improve on the face and content validity of the HASMID instrument. The original questionnaire was developed using a mixed methods approach, with items generated from interviews with PwD and an existing PRO measure [4]. Here we have been able to re-examine the content of the HASMID questionnaire through Patient and Public Involvement consultation and two focus group cognitive debriefing exercises. The alternatively phrased items were then subjected to psychometric evaluation alongside the existing HASMID items.

The development, refinement and evaluation of PRO instruments can be driven by different theoretical approaches: Classical Test Theory (CTT) Item Response Theory (IRT) and Rasch Measurement Theory (RMT) [19]. Here we applied CTT and RMT to assess the performance of the

alternatively worded items. Both assessments were used, in conjunction with clinical opinion to inform the final selection of items. This approach has been used by others in PRO development [20,21]. There are benefits to considering alternative sources of evidence. One of the main indicators for the inclusion of Rasch analysis in this study was to identify whether items demonstrated DIF. There is logic to considering eliminating items based upon differing responses driven by gender and age. However, in the context of self-management of diabetes, age and gender may well be factors that drive individuals' responses. For example, being tied to mealtimes may be a negative issue for the younger generation, whereas older people may already follow a more structured routine. The benefit of including psychometric analyses, Rasch analyses and clinical opinion allows a consideration of the relevance of the inclusion/exclusion of each item. Adopting only one approach may result in an instrument that is not relevant to the target population or provide information that is useful in trials, service evaluations, or routine clinical care.

The alternative items to the original HASMID items tested in this study were felt to be exploring the same concept, but with slightly different phrasing. For three (of the five) alternatively phrased items there were insufficient evidence to suggest an amendment to the wording of the original HASMID item. Two of the alternative items have now been incorporated into the HASMID questionnaire – Irritable and Social Activities. Both items performed well psychometrically, were deemed of clinical importance and captured different severity of underlying health to the original items.

One of the limitations of the study was that respondents self-reported information about their diabetes including diabetes type, duration of diabetes, HbA1c status, treatment and diabetes-related complications. Our HbA1c data indicates that respondents often did not know their HbA1c status, meaning that this cannot be reliably used to assess how HASMID performs across different levels of HbA1c, and we have no objective measure of severity by which to examine the items.

Responsiveness was examined; however the analysis is limited by the lack of an objective measure indicating change. The survey contains entirely self-report measures and whilst these can be used to indicate change, it should be recognised that these measures may not be fully accurate as they are self-reported by the respondent. Equally the study design where the survey is not attached to a clinical intervention means that there are no expectations for change in health or self-management for any particular individual. The definition of what constitutes “responsiveness” is not agreed, and this is a limitation for the summary of responsiveness results reported in table X. Approaches have been described on how to examine responsiveness and minimal important differences [22], however the figure of what constitutes a significant level of responsiveness is arbitrary. The same can also be said for floor and ceiling effects.

A potential limitation to the study is the applicability and performance of items across different ethnic groups. Within the psychometric survey information on ethnicity was not collected. This was a purposeful omission. It was outside the scope of the study to assess the cultural validity of the overall questionnaire and all potential items. To do so would require further qualitative work across different ethnic groups, specifically on cognitive debriefing to ensure items are relevant, and to identify any potential new items for consideration. Therefore it was felt that to collect data on the ethnicity of survey respondents would not be relevant. Any difference in HASMID scores across groups could not be validated at this stage. Further research is required to assess the cross-cultural validity of the HASMID-10 amongst different populations.

CONCLUSION

This cross-sectional validation study has examined the psychometric properties of the original HASMID items and tested potential items for inclusion using a large dataset. Rasch analysis was undertaken and considered alongside conventional psychometric performance. The analyses found the items to have good psychometric performance, with discriminative validity to be able to discriminate across

type of diabetes and treatment, and responsiveness to self-reported change in general health and general self-management. The final HASMID questionnaire now consists of ten items, the HASMID-10. The additional two items may provide further insight into how PwD are self-managing their condition, by providing further detail into how emotions and daily activities are affected.

Declarations

Ethics approval and consent to participate

The study was approved by the National Health Service Research Ethics Committee (NRES Committee West Midlands – Coventry and Warwickshire, Ref: 16/WM/03/45) and followed the tenets of the Declaration of Helsinki. All institutional research governance requirements were satisfied prior to any data collection.

Consent for publication

Not applicable.

Availability of data and material

The datasets during and/or analysed during the current study available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The study was funded by The Health Foundation.

Authors' contributions

JC, DR and JE designed the study. JC led the data collection. JC and DR performed data analyses. JC, DR and JE made substantial contributions to data interpretation. JC drafted the final manuscript. DR and JE made significant input to the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the Lay ADvice on Diabetes and Endocrine Research panel (LADDER) based at Sheffield Teaching Hospitals NHS Trust for their support and guidance.

References

1. National Diabetes Audit, 2015-16 Report 2a: Complications and mortality (complications of diabetes) England and Wales. 13th July 2017.
2. DAFNE Study Group. Training in flexible, intensive insulin management to enable dietary freedom in people with type 1 diabetes: dose adjustment for normal eating (DAFNE) randomised controlled trial. *BMJ* 2002; (7367):746.
3. National Clinical Guideline NG17. Methods, evidence and recommendations. Type 1 diabetes in adults: diagnosis and management. August 2015.
4. Carlton J, Elliott J, Rowen D, Stevens K, Basarir H, Meadows K, Brazier J. Developing a questionnaire to determine the impact of self-management in diabetes: giving people with diabetes a voice. *Health and Quality of Life Outcomes* 2017; 15.
5. Rowen DL, Labeit A, Stevens K, Elliott J, Mulhern B, Carlton J, Basarir H, Brazier J. Estimating a preference-based single index measuring the quality of life impact of self-management for diabetes. *Med Decis Making* 2018. doi: 10.1177/0272989X18784291. [Epub ahead of print]
6. Rowen DL, Stevens K, Labeit A, Elliott J, Mulhern B, Carlton J, Basarir H, Ratcliffe J & Brazier J. Using a Discrete Choice Experiment Involving Cost to Value a Classification System Measuring the Quality-of-Life Impact of Self-Management for Diabetes. *Value in Health* 2018; 21(2):69-77.
7. Lim CR, Harris K, Dawson J, et al. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open* 2015; doi: 10.1136/bmjopen-2015-007765
8. Marx RG, Jones EC, Atwan NC et al. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. *J Bone Joint Surg Am* 2005; 87(9-I):1999-2005
9. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995; 4:293-307

10. Wamper KE, Sierevelt IN, Poolman RW, et al. The Harris hip score: do ceiling effects limit its usefulness in orthopedics? *Acta Orthop* 2010; 81:703-7.
11. Terwee C, Bot MR, Bot M et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60:34-42.
12. Cohen J. A power primer. *Psychol Bull* 1992; 112(1):155-9.
13. Jacobson AM, de Groot M. The evaluation of two measure of quality of life in patients with type 1 and type 2 diabetes. *Diabetes Care* 1994; 14(4): 267-74.
14. Rasch G. *Probabilistic models for some intelligence and attainment tests*. 2nd ed. Chicago: University of Chicago Press; 1980.
15. Young T, Yang Y, Brazier JE, Tsuchiya A, Coyne K. The first stage of developing preference-based measure: constructing a health-state classification using Rasch analysis. *Qual Life Res* 2009; 18(2):253-65.
16. Young TA, Yang Y, Brazier JE, Tsuchiya A. The use of Rasch analysis in reducing a large condition specific instrument for preference valuation: The case of moving from AQLQ to AQL-5D. *Med Decis Making* 2011; 31(1):195-210.
17. Mulhern B, Rowen D, Jacoby A, Marson T, Snape D, Hughes D et al. The development of a QALY measure for epilepsy: NEWQOL-6D. *Epilepsy Behav* 2012;24(1):36-43.
18. Mulhern B, Smith SC, Rowen D, Brazier JE, Knapp M, Lamping DL et al. Improving the measurement of QALYs in dementia: developing patient- and carer-reported health state classification systems using Rasch analysis. *Value Health* 2012; 15(2):323-33.
19. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using classical test theory, item response theory and Rasch measurement theory to evaluate patient-reported outcome measures: A comparison of worked examples. *Value Health* 2015; 25-34.
20. Keetharuth AD, Taylor Buck E, Acquadro C, et al. Integrating qualitative and quantitative data in the development of outcome measures: the case of the Recovering Quality of Life (ReQoL)

measures in mental health populations. *Int. J. Environ. Res. Public Health* 2018; 15(7),1342.
doi.org/10.3390/ijerph15071342

21. Jin X, Liu GG, Gerstein HC, et al. Item reduction and validation of the Chinese version of diabetes quality-of-life measure. *Health Qual Life Outcomes* 2018; 16:78. doi.org/10.1186/s12955-018-0905-z
22. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes* 2006; 4:70.

	Time point 1 (%)				Time point 2 (%)			
	Online		Postal		Online		Postal	
	T1DM (n=708)	T2DM (n=1200)	T1DM (n=87)	T2DM (n=789)	T1DM (n=202)	T2DM (n=323)	T1DM (n=62)	T2DM (n=636)
HASMID ^{v1} score (SD)	0.56 (0.19)	0.65 (0.19)	0.67 (0.18)	0.78 (0.17)	0.59 (0.19)	0.69 (0.18)	0.72 (0.16)	0.78 (0.16)
EQ-5D-5L score (SD)	0.79 (0.23)	0.74 (0.27)	0.80 (0.24)	0.79 (0.24)	0.82 (0.20)	0.78 (0.24)	0.83 (0.21)	0.79 (0.24)

Table 2 Responsiveness and test-retest of items by general health and self-management for each sample

	T1DM				T2DM			
	General Health		Self-Management		General Health		Self-Management	
	Resp.	Test-retest	Resp.	Test-retest	Resp.	Test-retest	Resp.	Test-retest
Temper	‡	×	‡	×	×	✓	×	✓
Irritable	×	✓	×	✓	×	✓	×	✓
Hypo	†	✓	×	✓	×	✓	‡	✓
Tired	✓	✓	×	✓	✓	✓	×	✓
Tied Mealtimes	×	✓	†	✓	×	✓	×	✓
Affects Mealtimes	×	✓	×	✓	×	✓	×	✓
Control	‡	✓	✓	✓	×	✓	✓	✓
Daily Routine	✓	✓	‡	✓	†	✓	×	✓
Social Activities	×	✓	×	×	†	✓	×	✓
Hassle	‡	✓	‡	✓	×	✓	×	✓
Problem	‡	✓	✓	✓	×	✓	×	✓
Stressful	‡	✓	‡	×	†	✓	†	✓

Support	‡	✓	✘	✘	‡	✓	✓	✓
---------	---	---	---	---	---	---	---	---

Responsiveness ✓ = mean change from time point 1 to time point 2 >0.1

Responsiveness ✘ = mean change from time point 1 to time point 2 <0.1

Responsiveness † = mean change from time point 1 to time point 2 >0.1 for deterioration in general health/self-management

Responsiveness ‡ = mean change from time point 1 to time point 2 >0.1 for improvement in general health/self-management

Test-retest ✓ = mean change from time point 1 to time point 2 <0.1

Test-retest ✘ = mean change from time point 1 to time point 2 >0.1

Table 3 Floor and ceiling effects of items for each sample at both time points

	Postal Time Point 1				Postal Time Point 2				Online Time Point 1				Online Time Point 2			
	T1DM		T2DM		T1DM		T2DM		T1DM		T2DM		T1DM		T2DM	
	FE	CE	FE	CE												
Temper	x	✓	x	✓	x	✓	x	✓	x	x	x	✓	x	✓	x	✓
Irritable	x	x	x	x	x	x	x	x	✓	✓	x	x	x	x	x	x
Hypo	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	✓	✓	x	✓
Tired	✓	x	x	x	✓	x	x	x	✓	x	✓	x	x	x	✓	x
Tied Mealtimes	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓
Affects Mealtimes	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓
Control	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓
Daily Routine	x	✓	x	✓	x	✓	x	✓	x	x	x	✓	x	x	x	✓
Social Activities	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓	x	✓
Hassle	x	x	x	✓	x	x	x	✓	✓	x	x	✓	✓	x	x	✓
Problem	x	✓	x	✓	x	✓	x	✓	x	x	x	✓	x	x	x	✓
Stressful	x	✓	x	✓	x	✓	x	✓	✓	x	x	✓	x	x	x	✓

Support	✘	✓	✘	✓	✘	✓	✘	✓	✘	✘	✘	✘	✘	✘	✘	✓
---------	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

T1DM = Type 1 Diabetes Mellitus

T2DM = Type 2 Diabetes Mellitus

FE = Floor effects

CE = Ceiling Effects

✓ = significant i.e. >15%

✘ = non-significant i.e. <15%

Table 4 Summary of Rasch analysis

	Postal Time Point 1				Postal Time Point 2				Online Time Point 1				Online Time Point 2			
	DIF gender	DIF age	DIF DM Type	Disordere d Threshold s												
Temper	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗
Irritable	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗
Hypo	✗	✗	✓	✓	✓	✗	✓	✓	✗	✗	✓	✓	✗	✗	✓	✓
Tired	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✓	✗
Tied	✗	✓	✗	✗	✗	✓	✗	✗	✗	✓	✗	✗	✓	✓	✓	✗
Mealtime s																
Affects Mealtime s	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗

Control	✓	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Daily Routine	✗	✓	✓	✗	✗	✓	✓	✓	✗	✗	✓	✓	✗	✓	✓	✗
Social Activities	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗
Hassle	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✗
Problem	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✗
Stressful	✗	✓	✗	✗	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗
Support	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✓	✗	✗

Figure 1 HASMID^{v1} questionnaire

Dimension	Score	Wording
Mood	3	You <u>never</u> find yourself losing your temper over small things
	2	You <u>sometimes</u> find yourself losing your temper over small things
	1	You <u>usually</u> find yourself losing your temper over small things
	0	You <u>always</u> find yourself losing your temper over small things
Hypoglycaemic attacks	3	You <u>never</u> worry about going hypo
	2	You <u>sometimes</u> worry about going hypo
	1	You <u>usually</u> worry about going hypo
	0	You <u>always</u> worry about going hypo
Vitality	3	You are <u>never</u> tired
	2	You are <u>sometimes</u> tired
	1	You are <u>usually</u> tired
	0	You are <u>always</u> tired
Social Limitations	3	Your days are <u>never</u> tied to meal times
	2	Your days are <u>sometimes</u> tied to meal times
	1	Your days are <u>usually</u> tied to meal times
	0	Your days are <u>always</u> tied to meal times
Control	3	You feel you have <u>a lot of control</u> of your diabetes
	2	You feel you have <u>some control</u> of your diabetes
	1	You feel you have <u>little control</u> of your diabetes
	0	You feel you have <u>no control</u> of your diabetes
Hassle	3	You find your life with diabetes is <u>never</u> a hassle
	2	You find your life with diabetes is <u>sometimes</u> a hassle
	1	You find your life with diabetes is <u>often</u> a hassle
	0	You find your life with diabetes is <u>always</u> a hassle
Stress	3	You find your life with diabetes is <u>never</u> stressful
	2	You find your life with diabetes is <u>sometimes</u> stressful
	1	You find your life with diabetes is <u>often</u> stressful
	0	You find your life with diabetes is <u>always</u> stressful
Support (All support you have; from family, friends)	3	You feel <u>totally supported</u> with your diabetes
	2	You feel you have <u>a lot of support</u> with your diabetes
	1	You feel you have <u>a little support</u> with your diabetes

and health care professionals)	0	You feel you have <u>no support</u> with your diabetes
Additional items generated from Phase 1		
Mood	3	You are <u>never</u> irritable
	2	You are <u>sometimes</u> irritable
	1	You are <u>usually</u> irritable
	0	You are <u>always</u> irritable
Social Limitations	3	Your diabetes <u>never</u> affects your meal times
	2	Your diabetes <u>sometimes</u> affects your meal times
	1	Your diabetes <u>usually</u> affects your meal times
	0	Your diabetes <u>always</u> affects your meal times
Social Limitations	3	Your diabetes <u>never</u> affects your daily routine
	2	Your diabetes <u>sometimes</u> affects your daily routine
	1	Your diabetes <u>usually</u> affects your daily routine
	0	Your diabetes <u>always</u> affects your daily routine
Social Limitations	3	Your diabetes <u>never</u> limits your social activities
	2	Your diabetes <u>sometimes</u> limits your social activities
	1	Your diabetes <u>usually</u> limits your social activities
	0	Your diabetes <u>always</u> limits your social activities
Hassle	3	Your diabetes <u>never</u> causes you a problem
	2	Your diabetes <u>sometimes</u> causes you a problem
	1	Your diabetes <u>usually</u> causes you a problem
	0	Your diabetes <u>always</u> causes you a problem

HASMID-10

Mood	You <u>never</u> find yourself losing your temper over small things
	You <u>sometimes</u> find yourself losing your temper over small things
	You <u>usually</u> find yourself losing your temper over small things
	You <u>always</u> find yourself losing your temper over small things
Mood	You are <u>never</u> irritable
	You are <u>sometimes</u> irritable
	You are <u>usually</u> irritable
	You are <u>always</u> irritable
Hypoglycaemic attacks	You <u>never</u> worry about going hypo
	You <u>sometimes</u> worry about going hypo
	You <u>usually</u> worry about going hypo
	You <u>always</u> worry about going hypo
Vitality	You are <u>never</u> tired
	You are <u>sometimes</u> tired
	You are <u>usually</u> tired
	You are <u>always</u> tired
Social Limitations	Your days are <u>never</u> tied to meal times
	Your days are <u>sometimes</u> tied to meal times
	Your days are <u>usually</u> tied to meal times
	Your days are <u>always</u> tied to meal times
Social Limitations	Your diabetes <u>never</u> limits your social activities
	Your diabetes <u>sometimes</u> limits your social activities
	Your diabetes <u>usually</u> limits your social activities
	Your diabetes <u>always</u> limits your social activities
Control	You feel you have <u>a lot of control</u> of your diabetes
	You feel you have <u>some control</u> of your diabetes

	You feel you have <u>little control</u> of your diabetes
	You feel you have <u>no control</u> of your diabetes
Hassle	You find your life with diabetes is <u>never</u> a hassle
	You find your life with diabetes is <u>sometimes</u> a hassle
	You find your life with diabetes is <u>often</u> a hassle
	You find your life with diabetes is <u>always</u> a hassle
Stress	You find your life with diabetes is <u>never</u> stressful
	You find your life with diabetes is <u>sometimes</u> stressful
	You find your life with diabetes is <u>often</u> stressful
	You find your life with diabetes is <u>always</u> stressful
Support	You feel <u>totally supported</u> with your diabetes
	You feel you have <u>a lot of support</u> with your diabetes
	You feel you have <u>a little support</u> with your diabetes
	You feel you have <u>no support</u> with your diabetes