BMJ Open

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Marie Hauge Pedersen [1], Polly Duncan [2], Mathias Lasgaard,[1] Karina Friis,[1] Chris Salisbury [2], Finn Breinholt Larsen [1]

To cite: Pedersen MH, Duncan P, Lasgaard M, *et al*. Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study. *BMJ Open* 2022;**12**:e055276. doi:10.1136/bmjopen-2021-055276

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-055276).

Received 07 July 2021
Accepted 28 October 2021

[1]Public Health and Health Services Research, DEFACTUM, Central Denmark Region, Aarhus, Denmark
[2]Centre for Academic Primary Care, NIHR School for Primary Care Research, University of Bristol, Bristol, UK

Correspondence to
Dr Marie Hauge Pedersen; marie.pedersen@rm.dk

ABSTRACT

Objective To validate the Danish Multimorbidity Treatment Burden Questionnaire (MTBQ) and obtain a population-based evaluation of treatment burden.

Design Mixed-methods.

Setting Danish population-based survey.

Participants Translation by professional translators and an expert group. The scale was tested by 13 407 participants (aged ≥25 years) in treatment.

Measures The 10-item MTBQ was translated into Danish using forward-backward translation and used in a large population health survey. A global MTBQ score was calculated and factor analysis and Cronbach's alpha assessed dimensional structure and internal consistency reliability, respectively. Spearman's rank correlations between global MTBQ scores and scores of self-rated health, health-related quality of life and the number of long-term conditions, respectively, assessed construct validity. MTBQ scores were grouped into four categories (no, low, medium, high burden) to assess interpretability and population-based evaluation of treatment burden.

Results The scale showed high internal consistency (α=0.87), positive skewness and large floor effects. Factor analysis supported a one-dimensional structure of the scale with a three-dimensional structure as a less parsimonious alternative. The MTBQ score was negatively associated with self-rated health (r_s–0.45, p<0.0001) and health-related quality of life (r_s–0.46/–0.51, p<0.0001), and positively associated with the number of long-term conditions (r_s 0.26, p<0.0001) and perceived stress (r_s 0.44, p<0.0001). Higher treatment burden was associated with young age, male sex, high educational level, unemployment, being permanently out of work, not living with a spouse/cohabitant, living with child(ren) and long-term conditions (eg, heart attack, stroke, diabetes and mental illness).

Conclusion The Danish MTBQ is a valid measure of treatment burden with good construct validity and high internal reliability. This is the first study to explore treatment burden at a population level and provides important evidence to policy makers and clinicians about sociodemographic groups at risk of higher treatment burden.

Strengths and limitations of this study

- ► Using data from a large population health survey, we examined the validity of the Multimorbidity Treatment Burden Questionnaire (MTBQ) as a measure of treatment burden for identifying high-risk groups at the demographic level and guiding policy decisions and clinical practice.
- ► The response rate was high (64%), and weights were constructed to increase the generalisability of the analyses to the general population.
- ► A thorough process including forward-backward translation was undertaken to translate the MTBQ from English into Danish and to ensure the usability of the measure in a large population health survey.
- ► Establishment of content validity was out of scope for this paper and convergent validity was not established as the MTBQ was included in a comprehensive large-scale population survey, which precluded a comparative treatment burden measure.
- ► Responsiveness of the Danish MTBQ could not be tested due to the cross-sectional design.

INTRODUCTION

Treatment burden is defined as patients' perception of the effort required to look after their health and the effect of this on their everyday life.[1] Given the current disease-centred approach to healthcare, patients are often required to attend separate appointments, adopt lifestyle changes, self-monitor medical conditions and take complex combinations of medications. This can create considerable workload for patients, which may result in high treatment burden, particularly for patients with multimorbidity (multiple long-term conditions) and patients lacking the capacity or support to align their treatment with other roles and responsibilities in life.[2] A Danish ethnographical study suggested that patients with multimorbidity experienced difficulties in organising their treatment around their family life, social life and work life, and in setting goals and agendas with health professionals.[3]

Due in part to the world's ageing population, the prevalence of multimorbidity is increasing. It has now become the norm rather than the exception that patients presenting to general practice have multimorbidity, and healthcare systems are struggling to cope with the complexity.[4] As a result, there is a growing expectation that patients self-manage their health conditions,[5] which may lead to high treatment burden, potentially resulting in poor treatment compliance and low health-related quality of life.[6]

Evaluating the level of treatment burden in the general population may be used as a performance measure of how well the healthcare system responds to the population's needs while respecting their functioning and well-being.[7 8] Moreover, a study of treatment burden at the population level may identify subgroups at risk of poor outcomes who feel overwhelmed by their treatment and have problems with compliance, and therefore may benefit from less disruptive treatment.[9] To measure the burden of treatment at the population level and to better understand the relationship between treatment burden, demographic factors and health measures, a validated generic measure of treatment burden is essential.

In recent years, five generic measures of treatment burden (not specific to a particular health condition) have been developed[6 8 10–13] but none of these measures have yet been translated into Danish. We evaluated these five measures and we found the Multimorbidity Treatment Burden Questionnaire (MTBQ),[13] to be most suitable for use in a comprehensive population survey as it included key aspects of treatment burden and was concise with simple wording. The MTBQ is a 10-question measure of treatment burden developed and validated in a UK study of 1546 older adult participants (mean age 71 years) with three or more long-term conditions.[13] The MTBQ was validated as part of the 3D Study, a randomised controlled trial within primary care in England and Scotland.[13 14] In the UK, the scale has demonstrated good content and construct validity, reliability and responsiveness to change.[13] The aspects of treatment burden captured by the MTBQ includes medications, healthcare appointments, lifestyle changes, self-monitoring and having to rely on help from family and friends. These aspects are relevant to evaluate in a Danish context where the healthcare system is highly specialised with treatment regimens across specialities, sectors, and public and private healthcare providers, affecting many individuals with one or more conditions. We, therefore, considered the MTBQ a generic measure of treatment burden applicable to all persons in treatment although the scale was originally developed for individuals with multimorbidity.

The aims of this study were: (1) to translate the MTBQ into Danish; (2) to validate the psychometric properties of the Danish version of the MTBQ in a general population of individuals in treatment; and (3) to describe the relationship between treatment burden, sociodemographic factors and health measures at a population level. As far as we are aware, this is the first study to explore population-level associations between treatment burden and population characteristics.

METHODS

Study design and data collection

The study was based on cross-sectional data from the 2017 Danish population-based health survey named 'How are you?'. Denmark has approximately 5.7 million inhabitants and is divided into five administrative regions. This study comprises data from one of these regions: the Central Denmark Region. Approximately 23% of the Danish population resides in this region, which has a demographic composition similar to that of the total Danish population.[15] The survey consisted of a random sample drawn from the Danish Civil Registration System using the unique personal identification number assigned to all Danish residents. A total of 44 658 individuals aged 25 years and older were invited to participate in the survey (February–May 2017).

A total of 28 627 individuals (64%) responded to the questionnaire (web or postal). To identify participants in treatment, we used the following question as selection criteria: 'Do you receive treatment or take medication for one or more conditions, or do you attend rehabilitation or regular check-ups?' (yes/no). Respondents who reported to be in treatment were included in the study population and asked to complete the MTBQ. In total, 13 407 individuals in treatment comprised the study population for this study (online supplemental appendix A).

The MTBQ

In the original English MTBQ, individuals were asked how much difficulty they have with different aspects of treatment. Overall, the 10 questions covered the aspects of medication (three questions), self-monitoring (one question), contact with health professionals (three questions), obtaining information (one question), implementing lifestyle changes (one question) and relying on help (one question). Each question was scored from 0 (not difficult or does not apply), 1 (a little difficult), 2 (quite difficult), 3 (very difficult) to 4 (extremely difficult). A global score was calculated by taking an average of the questions answered and multiplying this by 25 (range 0–100). Participants were excluded if more than 50% of the answers were missing. To assess the interpretability of the questionnaire, the global MTBQ scores greater than 0 were grouped into tertiles resulting in four categories: no burden (score 0), low burden (score <10), medium burden (score 10–22) and high burden (score ≥22)).[13]

Sociodemographic factors

The following sociodemographic factors were included: sex, age, country of origin, deprivation level (register data), educational level, cohabitation status (living with a spouse/cohabitant or not), and living with child(ren) aged 0–15 years or not (survey data). Denmark was defined as country of origin if respondents had at least

one parent who was both a Danish citizen and was born in Denmark. Deprivation level was estimated for each of the 638 parishes in the Central Denmark Region based on the percentage with low educational level, the percentage people of working age without employment, and mean personal income. Using latent profile analysis,[16 17] we categorised parishes into five classes.[18] Parishes belonging to class one had the highest social ranking, whereas parishes belonging to class five were the most deprived. Respondents were asked about their highest level of completed school education and any further higher level of education. Using education nomenclature (International Standard Classification of Education) from Statistics Denmark, we categorised educational level as low (1–10 years of education), medium (11–14 years) and high (≥15 years). Employment status was defined by three categories: employed or student, unemployed (temporary or long term) and permanently out of work (disability pension, early retirement pension and old age pension) (combination of survey data and register data on receiving public benefit during the past year).[18]

Long-term conditions and multimorbidity

Information on long-term conditions was collected using a revised version of a disease checklist recommended by the WHO for health surveys.[19] Respondents were asked about 17 long-term and potentially fatal and/or debilitating conditions (table 1). They were recorded as having a condition if they currently had the condition or if they had previously had the condition and still experienced after-effects. When defining multimorbidity, we grouped some of the conditions together due to similarity in their risk factors or treatment regimens (table 1). The variable indicates the minimum number of long-term conditions as respondents may have had other conditions than these 17 specific conditions.

Measures of health status

From survey data, we collected information on self-rated general health (single question; five-point Likert scale), health-related quality of life (The Short Form (12) Health Survey (SF-12) V.2: Physical Component Summary and Mental Component Summary scores; calculated from 12 questions including the self-rated general health question; standardised (mean=50; SD=10) higher score indicated better health-related quality of life),[18 20] and perceived stress (Perceived Stress Scale 10-item version; ranging from 0 to 40; higher score indicated higher stress).[21 22]

Statistical analysis

Descriptive statistics of participant characteristics were generated for the study population. We tested the psychometric properties of the questionnaire against the minimum standards set out by the International Society for Quality of Life Research (ISOQOL).[23] The analysis plan and results are described in relation to ISOQOL's six recommended standards, and the Strengthening the

Table 1 Participant characteristics (participants in treatment aged 25+ years at the time of the 2017 'How are you?' survey, Central Denmark Region, n=13 407)

	n	%*
Sociodemographic factors		
Mean age (SD), years	59	16.0
Age, years (missing data: n=0; 0%)		
25–34	802	9
35–44	1368	12
45–54	2350	18
55–64	3142	21
65–74	3612	23
75–84	1720	13
85+	413	4
Sex (missing data: n=0; 0%)		
Female	7370	54
Male	6037	46
Country of origin (missing data: n=0; 0%)		
Denmark	12815	92
Other	592	8
Educational level (missing data: n=363; 2.7%)		
Low (0–10 years)	2387	21
Medium (11–14 years)	7336	54
High (15+ years)	3321	25
Employment status (missing data: n=161; 1.2%)		
Employed or student	5843	43
Unemployed	828	8
Permanently out of work	6575	49
Living with spouse/cohabitant (missing data: n=247; 1.8%)		
No	3318	32
Yes	9842	68
Living with child(ren) aged 0-15 years (missing data: n=1570; 11.7%)		
No	9815	82
Yes	2022	18
Deprivation level (missing data: n=0; 0%)		
1 Least deprived areas	1251	13
2	3784	28
3	3766	25
4	3966	27
5 Most deprived areas	640	8
Self-reported long-term conditions (missing data: n=57; 0.4%)		
Cardiovascular disease (one or more)	5471	39
Hypertension	5163	36
Angina pectoris	500	4
Heart attack	355	3
Stroke	488	4
Diabetes	1602	12
Cancer	876	6
COPD and/or asthma	2131	17

Continued

Table 1 Continued

	n	%*
COPD	1110	8
Asthma	1419	11
Allergy	2721	21
Osteoarthritis	4701	34
Rheumatoid arthritis	1358	11
Osteoporosis	1078	9
Slipped disk or other back disorder	2802	22
Mental illness	1982	17
Migraine or frequent headache	2273	18
Tinnitus	2372	17
Cataract	1065	8
No of self-reported long-term conditions† (missing data: n=57; 0.4%)		
0	1230	9
1	3403	26
2–3	5992	44
4+	2725	21
Mean no of self-reported long-term conditions (SD)	2.34	1.6
Scores of health measures		
Mean self-rated health score (SD)‡ (missing data: n=202; 1.5%)	2.98	0.9
Mean PCS score (SD)§ (missing data: n=0; 0%)	45.20	11.4
Mean MCS score (SD)§ (missing data: n=0; 0%)	47.39	11.1
Mean PSS score (SD) (missing data: n=169; 1.3%)	13.08	7.5

*Weighted to represent the population of the Central Denmark Region, aged 25+ years, in treatment.
†Number of self-reported conditions from a list of 17 conditions asked in the survey. When counting the number of conditions in an individual, conditions with similar risk factors and treatment regimens were grouped and only counted once. Hence, if a person had asthma and COPD, it was counted as one condition in that individual. Likewise, if a person had any combination of hypertension, heart attack and angina pectoris, it would only count as one condition in that individual. Respondents included in the analyses may have had other conditions than the 17 conditions asked in the survey.
‡Single question: 'In general, would you say your health is: excellent (5), very good (4), good (3), fair (2), poor (1)'?
§Based on the SF-12 V.2-questionnaire.
COPD, Chronic obstructive pulmonary disease; MCS, Mental Component Summary; PCS, Physical Component Summary; PSS, Perceived Stress Scale; SF-12, Short Form-12.

Reporting of Observational Studies in Epidemiology cross-sectional reporting guidelines were used.[24]

Wherever possible, the analyses were weighted to represent the population (25+ years) in treatment in the Central Denmark Region. The weights were constructed by Statistics Denmark to account for differences in selection probabilities and response rates between subgroups by linking respondents and non-respondents from the survey to registers and using a model-based calibration approach. Sensitivity analyses were conducted with a study sample restricted to respondents reporting minimum 2 of the 17 long-term conditions (n=8717). The latent profile analysis for deprivation level was conducted using Latent-GOLD, V.5. All other analyses were conducted using Stata, V.15.

Conceptual and measurement model

Conceptual framework, translation and pretest of the questionnaire

The original MTBQ was developed using the treatment burden framework developed by Eton *et al*.[1] A purpose of this study was to validate a Danish translated version of the MTBQ in a general population.

The translation from English into Danish were conducted in the following steps. Step 1: The MTBQ was forward translated into Danish by a professional native Danish registered translator, bilingual in English and Danish. Step 2: An expert group that comprised seven native Danish speakers, bilingual in English and Danish, reviewed the translation. The expert group had a background in questionnaire design, public health, health communication, anthropological fieldwork, multimorbidity research and nursing. Step 3: A blinded independent back translation was undertaken by a professional native English speaker, bilingual in Danish and English. Step 4: The expert group compared the backward version with the original English version.

Establishment of content validity was out of scope for this paper. However, a pretest was conducted to test the Danish translation of the MTBQ (including the purpose-maid treatment question) for inclusion in a large population survey, including the comprehensibility of the question formulations, and if the page layout was intuitive and functional. The pretest included two steps: First, 10 patients (native Danish speakers in a cardiac outpatient clinic; aged 18–90 years) were asked to fill out the questionnaire. Second, comprehensibility, meaningfulness, wording clarity and functionality of the layout was discussed through face-to-face interviews (online supplemental appendix B).

Question properties

Question properties were examined focusing on missing data, 'does not apply' responses and response distribution.

Dimensionality

In accordance with previous studies of treatment burden measures,[6 8 10–13] we used factor analysis to assess dimensionality. To test whether the one-dimensional structure of MTBQ discovered by Duncan *et al*[13] could be found in the Danish version of MTBQ, we estimated a confirmatory factor model (CFA)[25] with one factor using maximum likelihood with Satorra-Bentler (SB) adjustments to account for non-normal data.[26] Online supplemental analysis was conducted to investigate whether the fit of this model could be improved by including more factors. The specification of the number of factors, and which variables loaded onto which factor, was informed by exploratory factor analysis (details included in online

supplemental appendix C). As recommended by Hoyle and Panter,[27] model fit of CFA was evaluated with a range of fit indices including the χ^2, the root mean square error of approximation (RMSEA), the standardised root mean square residual (SRMR), the comparative fit index (CFI) and the Tucker-Lewis index (TLI). RMSEA ≤0.06 was interpreted as a close fit, while RMSEA ≤0.08 was interpreted as an acceptable fit.[28 29] A non-significant χ^2 and a value of SRMR ≤0.08 were taken as an acceptable fit, while for the CFI and the TLI a value of ≥0.95 was taken as an indicator of good fit.[29]

Reliability

Cronbach's α was calculated to assess internal consistency reliability with a score of 0.7–0.9 being deemed acceptable.[30]

Validity

Construct validity

The survey respondents rated each treatment aspect by level of difficulty (not difficult, a little difficult, quite difficult, very difficult, extremely difficult and does not apply). To ensure findings were comparable with other studies, the questions were scored as in the original UK validation study: from 0 (not difficult or does not apply) to 4 (extremely difficult). A global score was calculated by taking an average of the questions answered and multiplying this by 25 (range 0–100). Participants were excluded from the analysis if more than 50% of the answers were missing.

Based on the findings of prior studies, four prespecified hypotheses were tested to examine construct validity: first, a negative association between treatment burden and self-rated health[13]; second, a negative association between treatment burden and health-related quality of life[6 8 31]; third, a positive association between treatment burden and the number of self-reported long-term conditions[6 13 31 32]; and fourth, a positive association between treatment burden and perceived stress.[8] Spearman's rank correlation (r_S) was used to test the hypotheses because of the ordinal scoring of the measures, and Cohen's rule of thumb was used to interpret the magnitude of the associations (ie, r_S=|0.1| 'small'; r_S=|0.3| 'medium'; r_S=|0.5| 'large').[33]

Responsiveness

As the study was based on cross-sectional data, it was not possible to assess responsiveness to change.

Interpretability of scores

For the sake of comparison, we applied the same threshold scores as in the original UK study, that is, no burden (score 0), low burden (score <10), medium burden (score 10–22) and high burden (score ≥22). Across these categories, we compared participant characteristics (eg, age, sex, educational level, deprivation level, long-term conditions) and health measures (eg, self-rated health, health-related quality of life). Associations between treatment burden and participant characteristics were tested using logistic regression models that accounted for the ordered nature of the four category groupings of the MTBQ score. Standard ordered logit models, however, assume proportional odds, which empirically have been shown to be violated frequently.[33] We, therefore, used partial proportional odds models with which the proportional odds assumption could be tested using Wald tests and any possible violations could be adapted in the models.[34 35] When testing violations of the proportional odds assumption, we used a 1% significance level, as recommended, to minimise the significance of substantively trivial violations due to our large sample size.[34 35] We also estimated the models adjusted for age, sex, country of origin, educational level, employment status, marital status, living with child(ren) aged 0–15 years, deprivation level, individual long-term conditions and multimorbidity.

Demands on respondents

The effort required of respondents to complete the MTBQ questionnaire was assessed during the patient interviews in the pretest and by analysing the proportion of missing responses in the 'How are you?' survey.

Patient and public involvement

There was no patient or public involvement in the design of the study.

RESULTS

The participants were characterised by a mean age of 59 years, a small majority of women (54%), more than 9 out of 10 having a Danish origin, and more than half having a medium level of education (table 1). Almost half were permanently out of work and 43% were employed or studying. The large majority were living with a spouse or cohabitant (68%) and only a small proportion were living with child(ren) aged 0–15 years (18%). Around one-third were living in the most or second-most deprived areas and two-thirds had two or more long-term conditions. A global MTBQ score was calculated for 13 229 (99%) individuals.

Conceptual and measurement model

Conceptual framework, translation and pretest of the questionnaire

The MTBQ was translated into Danish in a thorough process that ensured that the meanings of the original questions were retained and easy to understand. During the process, the expert groups review of the translation resulted in changes to make the language slightly more informal. The comparison of the original English version and the backward translated English version showed good agreement and resulted in few minor changes of the Danish version. The pretest indicated that the questions were meaningful and easily understandable with clear formulations and satisfactory wording, but a clearer page layout of the 'Does not apply' response option was requested (online supplemental appendix B). The final Danish version was included in the 'How

Table 2 Responses to the Multimorbidity Treatment Burden Questionnaire (n=13 407)

Please tell us how much difficulty you have with the following:	N	Not difficult n (%*)	A little difficult n (%*)	Quite difficult n (%*)	Very difficult n (%*)	Extremely difficult n (%*)	Does not apply n (%*)	Missing data n (%)
1. Taking lots of medications	13 172	7840 (58)	1946 (15)	441 (4)	193 (2)	64 (1)	2688 (20)	235 (2)
2. Remembering how and when to take medication	13 209	9518 (70)	1468 (12)	260 (2)	149 (1)	70 (1)	1744 (14)	198 (1)
3. Collecting prescription medication	13 170	9775 (72)	1229 (10)	295 (3)	166 (2)	139 (1)	1566 (12)	237 (2)
4. Monitoring your medical conditions (eg, checking your blood pressure or blood sugar, monitoring your symptoms, etc)	13 112	7520 (55)	1409 (11)	343 (3)	156 (2)	113 (1)	3571 (28)	295 (2)
5. Arranging appointments with health professionals	13 150	8627 (64)	1350 (11)	379 (3)	159 (1)	137 (1)	2498 (19)	257 (2)
6. Seeing lots of different health professionals	13 083	6707 (49)	1512 (12)	486 (4)	239 (2)	180 (2)	3959 (30)	324 (2)
7. Attending appointments with health professionals (eg, getting time off work, arranging transport, etc)	13 088	7863 (58)	1439 (12)	436 (3)	256 (2)	151 (1)	2943 (23)	319 (2)
8. Obtaining clear and up-to-date information about your condition	13 084	8460 (62)	1710 (14)	551 (5)	262 (2)	185 (2)	1916 (15)	323 (2)
9. Making recommended lifestyle changes (eg, diet and exercise)	13 087	5494 (40)	2842 (21)	1224 (10)	563 (5)	375 (3)	2589 (20)	320 (2)
10. Having to rely on help from family and friends	13 134	4980 (37)	1398 (12)	606 (5)	335 (3)	278 (2)	5537 (41)	273 (2)

*Population-weighted proportions.

are you?'-questionnaire 2017, p.8, under the headline 'Behandling' ('Treatment').[36]

Question properties

The proportion of missing data for each question was less than 3% (table 2). Responses were positively skewed and high floor effects (60%–84%) were found for all questions. Furthermore, the global MTBQ score was positively skewed with 39% of participants scoring 0 (online supplemental appendix D). The highest proportion of 'does not apply' responses (41%) was seen for question 10.

Dimensionality

In the CFA analyses, the indices were ambiguous about the fit of the one-dimensional SB-adjusted model with no correlated errors (table 3). The χ^2 test was statistically significant, but it is well known that the test may be significant in large samples even though the magnitudes of model-data discrepancies are slight. This finding should, therefore, not lead to rejection of the model.[37 38] However, CFI (0.895) and TLI (0.865) did not indicate acceptable fit, whereas RMSEA (0.057) and SRMR (0.051) did indicate good fit. The standardised factor loadings ranged from 0.52 to 0.77 (all p<0.05) (online supplemental appendix E). Inspection of the modification indices revealed several correlated error terms, which is not that surprising since some of the questions covered the same underlying aspects (eg, use of medication). Allowing for a few correlated errors based on theoretical judgement markedly improved the model fit (table 3), supporting the unidimensionality of the MTBQ. In this case, the factor loadings ranged from 0.53 to 0.75 (all p<0.05).

In online supplemental appendix C, the exploratory factor analysis indicated the possibility of three factors. Based on this, a three-factor CFA was estimated. Factor 1, termed 'Medication and self-monitoring', included four items (nb. 1, 2, 3, 4), factor 2, termed 'Healthcare contacts and health information', included four items (nb. 5, 6, 7, 8) and factor 3, termed 'Coping ability', included two

Table 3 CFA models—global goodness-of-fit indices

	SB χ^2 (df)	P value	SB-RMSEA	SB-CFI	SB-TLI	SRMR
One-factor, with no correlated errors	1445.31 (35)	<0.0001	0.057	0.895	0.865	0.051
One-factor, with four correlations between errors*	680.22 (31)	<0.0001	0.041	0.952	0.930	0.036

Unweighted analyses.
*Correlations allowed between: (1) items 1 (taking lots of medications) and 2 (remembering how and when to take medication); (2) items 2 and 3 (collecting prescription medication); (3) items 5 (arranging appointments with health professionals) and 6 (seeing lots of different health professionals); (4) items 5 and 10 (Having to rely on help from family and friends).
CFI, Comparative Fit Index; RMSEA, the root mean square error of approximation; SB, Satorra-Bentler adjusted; SRMR, standardised root mean square residual; TLI, Tucker-Lewis Index.

items (nb. 9, 10). Overall, the results indicated that the three-factor CFA could be an alternative to the one-factor model with some improvements in model fit and small increases in factor loadings.

Reliability

The internal consistency of the scale was high with a population-weighted Cronbach's alpha of 0.87.

Validity

Construct validity

As expected, treatment burden correlated negatively with self-rated health (medium) and health-related quality of life (medium with physical health and large with mental health) and positively with the number of self-reported conditions (small) and perceived stress (medium) (table 4).

Responsiveness

Not applicable.

Table 4 Association between treatment burden score and self-rated health score, health-related quality of life score (PCS and MCS based on SF-12), number of self-reported conditions, and perceived stress score

Variable	N	Spearman's rank correlations (r_s)	P value
Self-rated health score*	13 032	**−0.45**	**<0.0001**
PCS score†	13 229	**−0.46**	**<0.0001**
MCS score†	13 229	**−0.51**	**<0.0001**
No of self-reported long-term conditions‡	13 180	**0.26**	**<0.0001**
PSS score	13 112	**0.44**	**<0.0001**

Bold values indicate statistical significance at the 5% level.
Unweighted analyses.
*Single question: 'In general, would you say your health is: excellent (5), very good (4), good (3), fair (2), poor (1)'?
†Based on the SF-12 V.2-questionnaire.
‡Number of self-reported long-term conditions from a list of 17 conditions in the survey, cf. table 1.
MCS, Mental Component Summary; PCS, Physical Component Summary; PSS, Perceived Stress Scale.

Interpretability of scores

Table 5 shows that the odds for perceiving higher treatment burden was highest among the youngest individuals (age 25–34 years) and decreased markedly until the age of 75–84 years after adjusting for sex, country of origin, educational level, employment status, cohabitation status, living with child(ren) aged 0–15 years, deprivation level, long-term conditions, and multimorbidity. The adjusted odds for higher treatment burden were also increased for males, individuals of Danish origin, highly educated individuals, unemployed individuals and those permanently out of work, individuals not living with a spouse or cohabitant, individuals living with child(ren) aged 0–15 years, individuals not living in the highest or lowest deprived areas, and for 13 specific long-term conditions (exceptions were asthma, allergy, tinnitus and cataract) with the highest odds seen among those affected by heart attack, stroke, diabetes and mental illness. Furthermore, the odds for higher treatment burden increased with the number of long-term conditions and with lower self-rated health, lower physical and mental health, and higher perceived stress. Six variables in the adjusted models did not fulfil the proportional odds assumption. For these variables, different ORs for no, low and medium treatment burden, respectively, were estimated (see table 5 and online supplemental appendix F for details).

Demands on respondents

During the translation process, great emphasis was put on ensuring that the wording was easy to understand and informal in a Danish context while retaining the original content of each question. In the test phase of the Danish version, none of the ten patients had difficulty understanding and answering the questions. Furthermore, the proportion of missing survey data for each question was less than 3% and a global MTBQ score was calculated for 99% of the participants, despite the fact that the 'How are you?' questionnaire is very comprehensive.

Sensitivity analyses

Overall, the sensitivity analyses supported the conclusions from the main analyses. Compared with the full sample, a small decrease was seen in the restricted sample (participants

Table 5 Participant characteristics by treatment burden categories (participants in treatment aged 25+ years at the time of the 2017 'How are you?' survey, Central Denmark Region, with a global MTBQ score)

		Treatment burden (score)*					
	N	None (0)	Low (<10)	Medium (10–22)	High (≥22)	Unadjusted OR (95% CI)†	Adjusted OR (95% CI)†‡
Participants (n, (%))	13229	5470 (39)	4066 (30)	2233 (18)	1460 (13)		
Sociodemographic factors							
Mean age (SD), years	13229	62 (15)	58 (16)	55 (16)	53 (16)		
Age, years (%)							
25–34 (ref.)	799	24	31	24	22		
35–44	1359	25	30	25	20	0.93 (0.78 to 1.12)	0.82 (0.68 to 1.00)
45–54	2333	31	32	20	17	**0.72 (0.61 to 0.85)**	**0.54 (0.45 to 0.66)**
55–64	3122	39	32	19	11	**0.49 (0.42 to 0.57)**	**0.34 (0.28 to 0.42)§**
65–74	3569	54	29	11	6	**0.26 (0.22 to 0.30)**	**0.15 (0.12 to 0.19)**
75–84	1661	52	27	12	8	**0.29 (0.24 to 0.35)**	**0.14 (0.10 to 0.18)**
85+	386	32	25	22	22	0.83 (0.63 to 1.08)	**0.41 (0.29 to 0.57)**
Sex (%)							
Female	7258	39	30	18	13	1.01 (0.94 to 1.09)	**0.84 (0.77 to 0.92)**
Male (ref.)	5971	39	30	18	13		
Country of origin (%)							
Denmark (ref.)	12647	40	31	18	12		
Other	582	33	24	20	23	**1.32 (1.08 to 1.61)§**	**0.75 (0.59 to 0.96)§**
Educational level (%)							
Low (0–10 years)	2315	39	28	18	16	**1.17 (1.05 to 1.29)**	1.11 (0.98 to 1.25)
Medium (11–14 years) (ref.)	7279	41	30	17	12		
High (15+ years)	3310	37	33	19	11	**1.19 (1.08 to 1.31)§**	**1.31 (1.17 to 1.47)§**
Employment status (%)							
Employed or student (ref.)	5820	37	32	19	11		
Unemployed	822	18	28	25	28	**2.78 (2.37 to 3.26)**	**1.58 (1.34 to 1.87)**
Permanently out of work	6441	45	28	15	12	**0.71 (0.65 to 0.77)§**	**1.22 (1.06 to 1.40)**
Living with spouse/cohabitant (%)							
No	3261	33	30	20	17	**1.59 (1.46 to 1.73)**	**1.44 (1.31 to 1.58)**
Yes (ref.)	9748	43	30	16	11		
Living with child(ren) aged 0–15 years (%)							
No (ref.)	9732	42	30	17	12		
Yes	2014	28	31	23	18	**1.76 (1.59 to 1.94)**	**1.16 (1.01 to 1.33)**
Deprivation level (%)							
1 Least deprived areas (ref.)	1234	43	31	17	9		
2	3737	39	32	17	12	**1.19 (1.05 to 1.35)**	**1.20 (1.04 to 1.39)**
3	3717	40	29	18	14	**1.26 (1.10 to 1.43)**	**1.21 (1.04 to 1.40)**
4	3917	39	30	18	14	**1.28 (1.12 to 1.45)**	**1.17 (1.01 to 1.36)**
5 Most deprived areas	624	32	28	22	18	**1.77 (1.46 to 2.14)**	1.13 (0.90 to 1.40)
Self-reported long-term conditions (%)¶							
Hypertension	5100	39	31	18	12	0.98 (0.91 to 1.06)	**1.27 (1.14 to 1.43)**
Angina pectoris	491	21	25	27	27	**2.67 (2.19 to 3.24)**	**1.77 (1.39 to 2.25)**
Heart attack	351	21	27	23	28	**2.54 (2.01 to 3.20)**	**2.05 (1.57 to 2.70)**
Stroke	480	26	22	21	31	**1.85 (1.46 to 2.35)§**	**2.21 (1.71 to 2.85)**
Diabetes	1580	26	32	23	19	**1.84 (1.65 to 2.06)**	**2.24 (1.94 to 2.58)**
Cancer	862	33	34	17	16	**1.25 (1.08 to 1.45)**	**1.59 (1.32 to 1.92)**

Continued

Table 5 Continued

	N	Treatment burden (score)* None (0)	Low (<10)	Medium (10–22)	High (≥22)	Unadjusted OR (95% CI)†	Adjusted OR (95% CI)†‡
COPD	1085	27	30	23	20	**1.82 (1.60 to 2.08)**	**1.74 (1.48 to 2.06)**
Asthma	1404	31	31	22	16	**1.45 (1.29 to 1.63)**	1.06 (0.92 to 1.23)
Allergy	2702	33	31	20	16	**1.40 (1.28 to 1.53)**	0.94 (0.82 to 1.06)
Osteoarthritis	4639	36	30	19	15	**1.25 (1.16 to 1.35)**	**1.24 (1.10 to 1.40)**
Rheumatoid arthritis	1340	28	27	22	22	**1.94 (1.70 to 2.20)**	**1.52 (1.30 to 1.78)**
Osteoporosis	1057	35	29	19	16	**1.25 (1.09 to 1.44)**	**1.35 (1.13 to 1.60)**
Slipped disk or other back disorder	2773	27	30	23	20	**2.01 (1.83 to 2.20)**	**1.53 (1.36 to 1.73)**
Mental illness	1964	16	26	27	30	**4.11 (3.70 to 4.56)**	**2.61 (2.27 to 3.00)**
Migraine or frequent headache	2252	25	27	24	24	**2.44 (2.20 to 2.70)**	**1.37 (1.20 to 1.57)**
Tinnitus	2342	37	29	19	15	**1.17 (1.06 to 1.29)**	1.02 (0.90 to 1.17)
Cataract	1045	39	28	17	16	1.06 (0.92 to 1.23)	1.11 (0.93 to 1.32)
No of self-reported long-term conditions (%)**							
0 (ref.)	1214	57	27	11	5		
1	3344	49	31	14	7	**1.36 (1.17 to 1.58)**	1.12 (0.93 to 1.35)
2–3	5936	38	31	18	13	**2.19 (1.90 to 2.52)**	**1.38 (1.09 to 1.76)**
4+	2686	22	28	25	25	**4.90 (4.19 to 5.72)**	**1.61 (1.10 to 2.35)**
Mean no of self-reported long-term conditions (SD)	13180	1.9 (1)	2.3 (2)	2.7 (2)	3.3 (2)		
Scores of health measures††							
Mean self-rated health score (SD)‡‡	13032	3.4 (1)	3.0 (1)	2.6 (1)	2.2 (1)	**0.38 (0.36 to 0.40)§**	**0.41 (0.39 to 0.44)§**
Mean PCS score (SD)§§	13229	51 (9)	46 (10)	41 (11)	34 (10)	**0.92 (0.92 to 0.93)§**	**0.91 (0.91 to 0.92)**
Mean MCS score (SD)§§	13229	53 (8)	48 (9)	42 (10)	35 (10)	**0.91 (0.90 to 0.91)§**	**0.92 (0.91 to 0.92)§**
Mean PSS score (SD)	13112	10 (6)	12 (7)	16 (7)	21 (7)	**1.12 (1.12 to 1.13)§**	**1.10 (1.09 to 1.11)§**

Bold values indicate statistical significance at the 5% level.
*Population-weighted means, SDs and proportions.
†Population-weighted partial proportional odds models.
‡Adjusted for age, sex, country of origin, educational level, employment status, marital status, living with child(ren) aged 0–15 years, deprivation level, individual long-term conditions and multimorbidity.
§The variable does not fulfil the proportional odds assumption in the ordered logit model at a 1% significance level. The regression coefficient (and OR) is, therefore, allowed to vary in the partial proportional odds model. The results can be seen in online supplemental appendix F.
¶The reference group for each self-reported long-term condition is the group of individuals without that specific condition. For example, for individuals with diabetes the reference group is those individuals without diabetes.
**Number of self-reported conditions from a list of 17 conditions asked in the survey. When counting the number of conditions in an individual, conditions with similar risk factors and treatment regimens were grouped and only counted once. Hence, if a person had asthma and COPD, it was counted as one condition in that individual. Likewise, if a person had any combination of hypertension, heart attack and angina pectoris, it would only count as one condition in that individual. Respondents included in the analyses may have had other conditions than the 17 conditions asked in the survey.
††The scores of health measures are included in the regression analyses as continuous variables. Hence, the ORs represent the odds of a higher level of treatment burden category versus the current or lower level treatment burden category for a one-unit increase in the health measure score.
‡‡Single question: 'In general, would you say your health is: excellent (5), very good (4), good (3), fair (2), Poor (1)'?
§§Based on the SF-12 V.2-questionnaire.
COPD, chronic obstructive pulmonary disease; MCS, Mental Component Summary; MTBQ, Multimorbidity Treatment Burden Questionnaire; PCS, Physical Component Summary; PSS, Perceived Stress Scale.

with ≥2 long-term conditions) in the proportion of 'Does not apply' responses (9%–35% vs 12%–41%), in the total floor effects (55%–82% vs 60%–84%) and in the global MTBQ score of 0 (37% vs 39%). In contrast to the original analysis, country of origin, living with child(ren) under the age of 16 and 4+ long-term conditions (reference: 2–3 long-term conditions) were not significantly associated with increased treatment burden, whereas the most deprived area was (reference: least deprived area).

DISCUSSION

In this study, we applied meticulous methods to translate and validate a Danish version of the MTBQ using data

from a large population health survey, which provided a population-based evaluation of treatment burden. The results indicate a high-quality Danish version that is easy to understand with satisfactory psychometric properties. The percentage of missing data for each question was small (<3%) despite the Danish survey being very comprehensive, but the floor effects of each question were high (60%–84%). There was evidence of high internal consistency and good construct validity and confirmatory factor analyses supported the unidimensionality of the scale, although online supplemental analysis indicated that a three-dimensional scale may be an alternative or complement to the one-dimensional scale.

At the population level, we found positive associations between treatment burden and a broad range of sociodemographic factors and health measures, including younger age, being male, being highly educated (compared with medium educational level), being unemployed or permanently out of work, not living with a spouse or cohabitant, living with child(ren), low self-rated health, poor health-related quality of life and high perceived stress. Also, most of the included long-term conditions were associated with higher treatment burden with the strongest associations being found among those affected by myocardial infarction, stroke, diabetes and mental illness; the odds for high treatment burden increased with the number of long-term conditions.

A key strength of our study is its use of a thorough translation process. This process included using the forward-backward translation method and pre-testing the translated version to ensure a high-quality translation. Another major strength was the use of a large population health survey with a high response rate, yielding a large study population, demonstrating that the MTBQ is suitable for evaluation of treatment burden in the general Danish population in treatment despite being developed for individuals with multimorbidity.[13] Moreover, the use of register-based information allowed for weighting of the data. Therefore, the findings of the study are likely to be representative of the general Danish population (aged 25+years) in treatment. A further strength of the study is that the MTBQ questionnaire was developed using the same conceptual framework[1] as the Patient Experience with Treatment and Self-Management (PETS) treatment burden measure.[8] Triangulation of findings across studies, cultures and healthcare systems strengthens the growing body of evidence about treatment burden and associated factors.

There are some limitations. First, to test construct validity, we tested prespecified hypotheses of associations between the MTBQ and selected health measures. Ideally, we would have included direct measures of treatment burden, such as number of tablets, complexity of medication regimen, and number of appointments with different health professionals. Additionally, it would have been preferable to include a comparative measure of treatment burden (eg, the Healthcare Task Difficulty[12] or the Treatment Burden Questionnaire[6]). This was precluded due to length limitations in the comprehensive population survey. Also, information on specific types of mental illness was not included in the survey data. Linking to register data on medical information could meet some of these limitations but this was out of scope for this study. Second, it was not possible to calculate the Spearman's rank correlations or to use the CFA models with SB adjustments using weighted data, which may reduce the generalisability of those findings. However, the weighted partial proportional odds models supported the associations between treatment burden and the health measures. Third, the responsiveness of the MTBQ was not assessed due to the data being cross-sectional, and therefore, no conclusions about temporality or causation can be made. Fourth, establishment of content validity was out of scope for this study. However, there is no reason to think that the range of topics which were important for patients in Denmark would be any different from those in the original validation study in the UK. Additionally, the pretest of the Danish MTBQ indicated good content validity.

In addition to the above limitations, there are some considerations regarding the original English MTBQ. First, as in other treatment burden measures[6 8 10 12] the rating scale is not balanced, but concentrates on the dimension of 'difficulty'. A balanced scale would include the dimension of 'easy' as well. This may affect the respondents' perception when answering the questionnaire. Second, in the scoring of the MTBQ the response options 'Not difficult' and 'Does not apply' was equated, potentially affecting the score distribution. Given the large floor effects, a revised approach to these considerations may be investigated in future research. Third, the applied threshold values for interpretability is based on a purely statistical criteria and not a clinical anchor, which would be relevant to identify in order to increase the clinical significance of the burden levels.

In accordance with our findings, previous studies have found that high treatment burden was associated with young age,[10 13] an increasing number of long-term conditions,[12 13 32] low health-related quality of life,[12 13] poor self-rated health,[13] mental illness,[11–13] stroke[12] and diabetes.[10] However, the reported positive association with male sex was not found in any of these previous studies. It is also noteworthy that Duncan *et al*[13] found no associations between high treatment burden and physical medical conditions, in contrast to this study. This difference may be due to our large sample size, giving sufficient power to detect significant associations even for less frequently occurring physical conditions. However, treatment regimens for specific conditions may also differ between the UK and Denmark, which may contribute to differences in the perceived treatment burden for those conditions. Additionally, in the UK study the conditions were obtained from medical record data, whereas for this study they were self-reported, which may contribute to the different findings. The positive association between treatment burden and high educational level (compared with medium level) is not a common finding of previous

studies either. This association may relate to high expectations among the highly educated to themselves, to the life they want to live and to treatment and the healthcare system.

High floor effects similar to those found in this study have previously been found in other studies of generic treatment burden measures.[10 12] However, both floor effects and positive skewness of the single questions and the global MTBQ score were more pronounced in this study than in the UK study.[13] One reason for this difference is that this study included participants taking treatment for one or more long-term conditions (mean age 59 years), whereas the UK study included a less well study population whereby participants had three or more long-term conditions (mean age 71 years). Sensitivity analyses showed a reduction in the 'Does not apply' answers and overall floor effects when the analysis was restricted to people with two or more long-term conditions, though the differences were small compared with the original findings. This indicates that the observed floor effects are primarily explained by respondents not feeling burdened by their treatment or experiencing that particular treatment task. It may not be surprising that many individuals do not have a problem with all included aspects. However, even having a large burden from one or two aspects can represent an important burden. Another possibility is that individuals most overwhelmed by their treatment may not have the capacity to fill out the comprehensive survey, thereby enhancing the skewed distribution of treatment burden if the weighting does not fully account for the possible response bias. Hence, the reported treatment burden distribution may be viewed as a conservative estimate of patient-perceived treatment burden in the general population. Also, even though high floor effects are common to treatment burden measures, the measures still show robust and meaningful relationships with other person-centred and patient-reported outcomes, supporting their usability. Nevertheless, different ways to address the issue of high floor effects will be explored in subsequent research.

When assessing dimensionality of a scale, optimisation of statistical fit should in general be balanced against parsimony where only factors that represent a considerable proportion of the variation in the data should be retained. The results from the main analyses support the possibility of using the Danish MTBQ as a one-dimensional measure of (overall) treatment burden across the different aspects of treatment burden. The advantage of this is a parsimonious description of the data, which has implications for use and interpretation. On balance we recommend a one-dimensional structure for the Danish MTBQ. We recognise, however, that there are advantages to having sub-dimensions, namely the ability to identify specific aspects of treatment that are particularly burdensome, which may be more difficult to identify when the MTBQ is used as an overall one-dimensional scale. We, therefore, recommend that subdomain analyses of specific aspects of treatment burden are explored in future research.

To the best of our knowledge, this is the first study to translate and validate a Danish version of a generic treatment burden measure. It is also the first study to assess sociodemographic factors and health measures associated with treatment burden at a population level, providing important evidence to policy makers and clinicians about groups of patients who are most likely to experience high treatment burden (eg, younger people, unemployed, individuals with childcare responsibilities). We recommend further validation to assess the responsiveness of the MTBQ in a Danish context and to assess the use of the MTBQ in a clinical setting, including identification of clinically relevant threshold values to identify patients with high, medium and low treatment burden, respectively, as well as establishing content validity. Also, subsequent research is planned to evaluate treatment burden using a more comprehensive list of long-term conditions drawing on medical register data. Additionally, research into the effects of different multimorbidity patterns on treatment burden is planned to strengthen the understanding of particular burdensome disease combinations for use in the planning of courses of treatment and healthcare system organisation.

Twitter Polly Duncan @polly_duncan and Chris Salisbury @prof_tweet

Acknowledgements The authors wish to thank the ten patients who took the time to participate in the pretest during the translation process as well as colleagues, among others Lone Flarup and Esben Dahl Nielsen, who contributed to the translation of the MTBQ and collection of survey data. We also wish to thank other colleagues of ours for valuable comments during the writing of this paper.

Contributors FBL was responsible for the translation process. MHP, PD, FBL, CS and KF participated in the design of the further validation and population-based evaluation with MHP being the main person responsible for the final design of this part. FBL, MHP, KF and ML were responsible for the collection of survey data. Data analyses were conducted by MHP with contributions from FBL, KF and ML. MHP drafted the scientific manuscript with contributions from PD and FBL. All authors contributed to interpretation, critically revised the paper, contributed to the final draft, approved the final manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. MHP is the guarantor.

Funding The 2017 'How are you?' survey was funded by the Central Denmark Region (Region Midtjylland).

Disclaimer The views and opinions expressed in the article are those of the authors and do not necessarily reflect those of the Central Denmark Region.

Editor's note Copyright The MTBQ was developed by Professor Chris Salisbury and Dr Polly Duncan. Copyright (including the Danish version) belongs to the University of Bristol but it is freely available for use under licence. Please see https://www.bristol.ac.uk/primaryhealthcare/resources/mtbq/ for details. Permission was obtained to translate the MTBQ into Danish.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The survey was approved by the Danish Data Protection Agency (r. no. 2012-58-0006) and registered in the Central Denmark Region (j. no: 1-16-02-593-16). Each participant received written information about the purpose of the survey, and informed consent was obtained from all subjects. All methods were carried out in accordance with relevant guidelines and regulations along with the approval.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been

ORCID iDs
Marie Hauge Pedersen http://orcid.org/0000-0001-8472-0555
Polly Duncan http://orcid.org/0000-0002-2244-3254
Chris Salisbury http://orcid.org/0000-0002-4378-3960
Finn Breinholt Larsen http://orcid.org/0000-0002-6848-123X

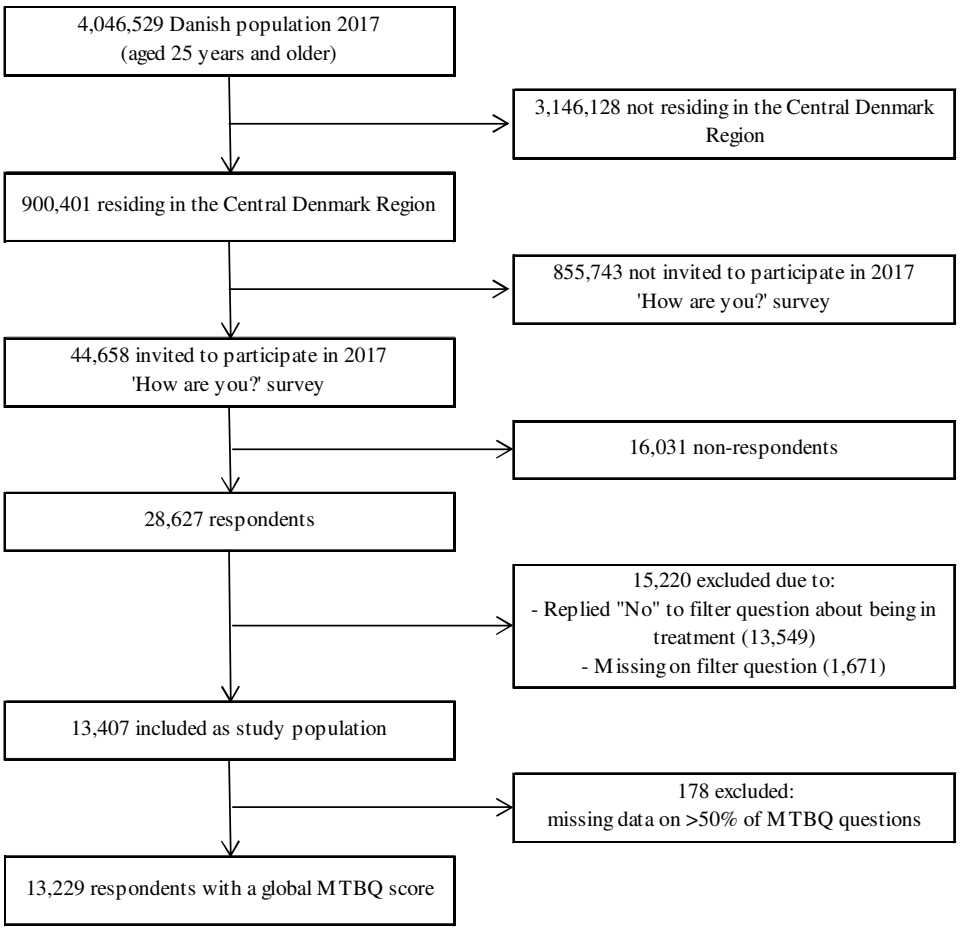
REFERENCES

1 Eton DT, Ramalho de Oliveira D, Egginton JS, *et al*. Building a measurement framework of burden of treatment in complex patients with chronic conditions: a qualitative study. *Patient Relat Outcome Meas* 2012;3:39–49.
2 Shippee ND, Shah ND, May CR, *et al*. Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice. *J Clin Epidemiol* 2012;65:1041–51.
3 Ørtenblad L, Meillier L, Jønsson AR. Multi-morbidity: a patient perspective on navigating the health care system and everyday life. *Chronic Illn* 2018;14:271–82.
4 Salisbury C, Johnson L, Purdy S, *et al*. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2011;61:e12–21.
5 May CR, Eton DT, Boehmer K, *et al*. Rethinking the patient: using burden of treatment theory to understand the changing dynamics of illness. *BMC Health Serv Res* 2014;14:281.
6 Tran V-T, Harrington M, Montori VM, *et al*. Adaptation and validation of the Treatment Burden Questionnaire (TBQ) in English using an Internet platform. *BMC Med* 2014;12:109.
7 Spencer-Bonilla G, Quiñones AR, Montori VM, *et al*. Assessing the burden of treatment. *J Gen Intern Med* 2017;32:1141–5.
8 Eton DT, Yost KJ, Lai J-S, *et al*. Development and validation of the Patient Experience with Treatment and Self-Management (PETS): a patient-reported measure of treatment burden. *Qual Life Res* 2017;26:489–503.
9 May C, Montori VM, Mair FS. We need minimally disruptive medicine. *BMJ* 2009;339:b2803.
10 Tran V-T, Montori VM, Eton DT, *et al*. Development and description of measurement properties of an instrument to assess treatment burden among patients with multiple chronic conditions. *BMC Med* 2012;10:68.
11 Gibbons CJ, Kenning C, Coventry PA, *et al*. Development of a multimorbidity illness perceptions scale (MULTIPleS). *PLoS One* 2013;8:e81852.
12 Boyd CM, Wolff JL, Giovannetti E, *et al*. Healthcare Task Difficulty among older adults with multimorbidity. *Med Care* 2014;52(Suppl 3):S118–25.
13 Duncan P, Murphy M, Man M-S, *et al*. Development and validation of the multimorbidity treatment burden questionnaire (MTBQ). *BMJ Open* 2018;8:e019413;8.
14 Salisbury C, Man M-S, Bower P, *et al*. Management of multimorbidity using a patient-centred care model: a pragmatic cluster-randomised trial of the 3D approach. *Lancet* 2018;392:41–50.
15 Statistics Denmark. FOLK1: population at the first day of the quarter by municipality, sex, age, marital status, ancestry, country of origin, and citizenship, 2018. Available: www.statistikbanken.dk/FOLK1 [Accessed 8 May 2018].
16 Bartholomew DJ, Knott M, Moustaki I. *Latent variable models and factor analysis: a unified approach*. West Sussex: John Wiley & Sons, 2011.
17 Bacher J, Vermunt JK. *Analyse latenter Klassen. Handbuch der sozialwissenschaftlichen Datenanalyse*. Springer, 2010: 553–74.
18 Larsen FB, Pedersen MH, Lasgaard M, *et al*. *Hvordan har du det? 2017 - Sundhedsprofil for region og kommuner (Bind 1). [How are you? 2017 - Health and morbidity profile for region and municipalities (Volume 1)]*. Aarhus: DEFACTUM, Region Midtjylland, 2018.
19 Buratta V, Frova L, Gargiulo L, *et al*. Development of a common instrument for chronic physical conditions. In: Nosikov A, Gudex C, eds. *EUROHIS: developing common instruments for health surveys*. IOS Press, 2003: 21–34.
20 Ware JE, Kosinski M, Turner-Bowker DM, *et al*. How to score version 2 of the SF-12 health survey (with a supplement documenting version 1). *Quality Metric* 2002.
21 Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;24:385–96.
22 Cohen S. *Perceived stress in a probability sample of the United States. The social psychology of health*. Thousand Oaks, CA, US: Sage Publications, Inc, 1988: 31–67.
23 Reeve BB, Wyrwich KW, Wu AW, *et al*. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905.
24 von Elm E, Altman DG, Egger M, *et al*. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61:344–9.
25 Kline R. Exploratory and confirmatory factor analysis. In: Petscher Y, Schatschneider C, Compton DL, eds. *Applied quantitative analysis in education and the social sciences*. New York, NY: Routledge, 2013: 183–217.
26 Satorra A, Bentler PM. *Corrections to test statistics and standard errors in covariance structure analysis. Latent variables analysis: applications for developmental research*. Thousand Oaks, CA, US: Sage Publications, Inc, 1994: 399–419.
27 Hoyle RH, Panter AT. Writing about structural equation models. In: Hoyle RH, ed. *Structural equation modeling: concepts, issues, and applications*. London: Sage, 1995: 158–98.
28 Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing structural equation models*. Newbury Park, CA: Sage Publications, 1993: 136–62.
29 Hu Li-tze, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:1–55.
30 Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. 4th ed. New York, NY: Oxford University Press, 2008.
31 Herzig L, Zeller A, Pasquier J, *et al*. Factors associated with patients' and GPs' assessment of the burden of treatment in multimorbid patients: a cross-sectional study in primary care. *BMC Fam Pract* 2019;20:88.
32 Sav A, Whitty JA, McMillan SS, *et al*. Treatment burden and chronic illness: who is at most risk? *Patient* 2016;9:559–69.
33 Long JS, Freese J. *Regression models for categorical dependent variables using stata*. 3rd edn. College Station, TX: Stata Press, 2014.
34 Williams R. Understanding and interpreting generalized ordered logit models. *J Math Sociol* 2016;40:7–20.
35 Williams R. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata J* 2006;6:58–82.
36 DEFACTUM. Hvordan har du det? En undersøgelse af trivsel, sundhed og sygdom blandt voksne i Region Midtjylland 2017 [How are you? 2017, Central Denmark Region questionnaire]. DEFACTUM - Region Midtjylland, 2017. Available: https://www.defactum.dk/siteassets/defactum/3-projektsite/hvordan-har-du-det/hhdd-2017/2017-voksen-sporgeskema_web.pdf [Accessed 10 Feb 2017].
37 Schmitt TA. Current methodological considerations in exploratory and confirmatory factor analysis. *J Psychoeduc Assess* 2011;29:304–21.
38 Tanaka JS. "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Dev* 1987;58:134–46.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix A

Appendix A: Selection of participants for the validation of the Danish version of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and population-based evaluation of treatment burden, 2017



Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix B

Appendix B: Pretest of the Danish MTBQ during the translation process

Objective:
The objective of the pretest was to investigate if the Danish translation of the MTBQ was suitable for inclusion in a large population survey, including whether the filter-question (developed to ensure that only respondents undergoing treatment would be asked to fill out the MTBQ) worked, whether the comprehensibility and meaningfulness of the question formulations was good, and whether the page layout was intuitive and functional.

Study design and data collection
The pretest was conducted in autumn 2016. Ten randomly selected patients and relatives aged 18-90 years, a mix of men and women, all native Danish speakers, with a majority of elderly, were recruited in a cardiac outpatient clinic waiting room at Aarhus University Hospital. Time constraints between the development of the final version of the English MTBQ and the deadline for the 'How are you?'-questionnaire left only little time for the recruitment process. Aarhus University Hospital was one of the largest and leading hospitals in Denmark (https://www.en.auh.dk/about-the-hospital/the-worlds-best-hospitals/) and the selected hospital department was known to have a heterogeneous patient composition comprising many patients with diverse combinations of multimorbidity but also patients with only one condition. Therefore, given the constraints for the process, this department was viewed as the best option as place of recruitment of a diverse set of patients.

Methods:
The pretest included two steps. First, the 10 participants were asked to fill out a questionnaire for themselves that comprised the filter-question and the Danish translated version of the MTBQ. Second, comprehensibility, meaningfulness, wording clarity, and functionality of the layout was discussed through face-to-face interviews based on the completed questionnaires. Among others, participants were asked if they understood the filter-question and if this made sense to them and they were asked about each specific question formulation; how easy it was to understand it, whether it was meaningful, and whether the language / word choices were appropriate.

Results:
Regarding the filter-question
Two young participants under the age of 20 answered "No" in the filter question, while the other (older) participants answered "Yes". The participants' answers to the filter-question along with the subsequent dialogue on this indicated that the question would be useful as a filter in the large population survey to ensure that only persons in treatment for long-term conditions would complete the MTBQ.

Regarding the MTBQ
The eight people who had answered "Yes" in the filter question and who therefore "went on" to the MTBQ filled out the entire questionnaire.

One participant commented that she had not seen the "Does not apply" response option until after she had filled out the questionnaire. Apart from that, the participants only expressed that the questions were finely worded with a well-chosen language that made them easy to understand and worked as intended. The participants made no further proposals for changes.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix B

The total duration was about four hours.

Conclusion

All in all, the results from the pretest indicated that the questions in the tested version were well translated so that they appeared well-formulated, meaningful, and with satisfactory wording to be well-functioning and easily understandable in a broad Danish context. The only change resulting from the pretest was a clearer page layout of the "Does not apply" response option.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix C

Appendix C: Supplementary factor analysis

As a supplementary analysis, we conducted a confirmatory factor analysis (CFA) with more than one factor. This was to investigate if a more complex model would provide a markedly better fit to the data compared to the initial one-factor model that was investigated and found support of in the main analyses.

The specification of the CFA model was informed by an exploratory factor analysis (EFA) (principal factor; PF). EFA was used to investigate whether there was indication of more than one possible dimension in the Danish MTBQ and to guide the partitioning of the ten items comprised by the MTBQ into different dimensions.

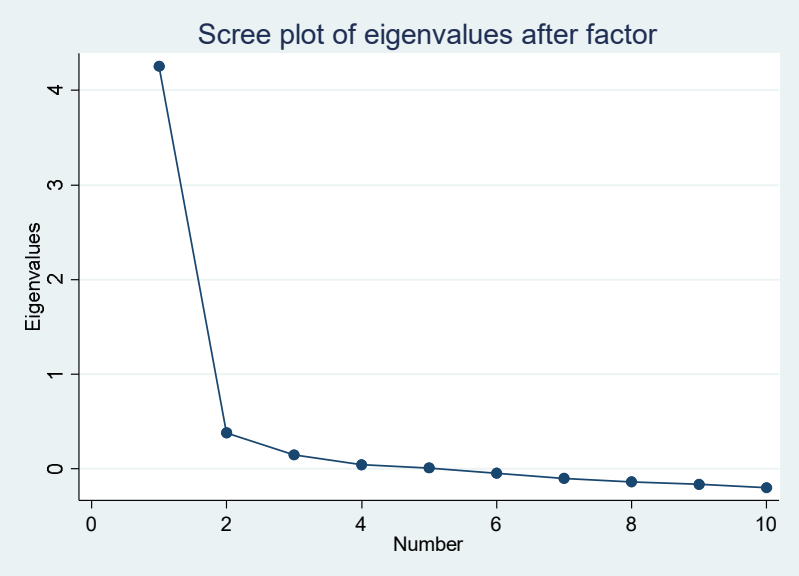
The average of the communalities in the PF analysis was 0.42. Only one factor (Factor 1) had an eigenvalue larger than this average and for all items the highest loading was seen on this factor (uniformly > 0.4), supporting a one-factor model (table 1).

Table 1. Principal factor (PF) analysis

Factor	Eigenvalue	Factor loading range
Factor 1	4.25503	0.55 to 0.76
Factor 2	0.38289	-0.25 to 0.31
Factor 3	0.14953	-0.14 to 0.22
Factor 4	0.04474	-0.08 to 0.12
Factor 5	0.01202	-0.04 to 0.07

Furthermore, the scree plot clearly indicated one primary explanatory factor (figure 1).

Figure 1. Scree plot



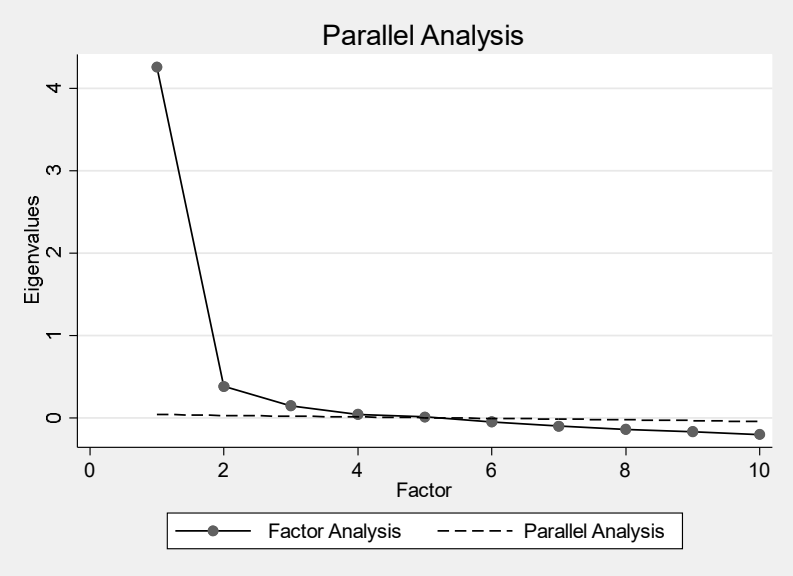
Based on factor loadings, eigenvalues and scree plot, the one-factor model could be argued to be favoured. These results do not indicate important neglected factors with high explanatory power, though the scree plot may indicate the presence of one or two additional factors.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix C

A parallel analysis (PA) was also conducted. The recommendation is to retain factors where the observed eigenvalue from the factor analysis (FA) is larger than the simulated eigenvalue from the PA analysis. Based on this criteria, a PA with 1,000 iterations suggested that five factors were retained for rotation of the factor loadings (figure 2 and table 2), though the difference in eigenvalues in factor 4 and 5 are quite small.

Figure 2. Parallel analysis
N = 12,370



PA eigenvalues averaged over 1000 replications.

Table 2. Parallel analysis - eigenvalues
N = 12,370

	FA	PA	Dif
1	4.255026	0.044833	4.210194
2	0.382889	0.032239	0.35065
3	0.149526	0.022284	0.127242
4	0.044736	0.013131	0.031604
5	0.012015	0.004554	0.007461
6	-0.04511	-0.0037	-0.04141
7	-0.09914	-0.01201	-0.08713
8	-0.13606	-0.02082	-0.11524
9	-0.16362	-0.03056	-0.13306
10	-0.19796	-0.04265	-0.15532

PA eigenvalues averaged over 1000 replications.
Dif, difference in eigenvalues; FA, factor analysis; PA, parallel analysis.
.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix C

Oblique rotation was used to allow for correlation between the factors. The rotated factors are presented in table 3.

Table 3. Rotated factors (oblique rotation) after FA of the Danish MTBQ

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Uniqueness
Item 1	0.0351	0.0986	**0.5695**	-0.0196	-0.0161	0.5798
Item 2	0.0469	0.0022	**0.6293**	0.0331	0.0552	0.4998
Item 3	0.1481	0.0219	**0.3862**	0.0068	0.2558	0.5412
Item 4	0.2733	0.0304	**0.3479**	0.152	0.0046	0.5176
Item 5	**0.7561**	-0.0913	0.0588	0.0541	0.0617	0.3559
Item 6	**0.7393**	0.0773	0.0508	-0.0175	-0.0369	0.3607
Item 7	**0.5060**	0.1096	0.0094	-0.0221	0.1672	0.5294
Item 8	**0.5321**	0.2264	0.0388	0.0047	-0.055	0.4872
Item 9	0.0752	**0.445**	0.0509	0.1065	-0.0246	0.6562
Item 10	0.0499	**0.4801**	0.1022	-0.0261	0.0644	0.6280

For each item the highest loading is marked with bold.

The highest loadings on the rotated factors were distributed on three factors:

"Medication and self-monitoring": item 1 (taking lots of medications), item 2 (remembering how and when to take medication), item 3 (collecting prescription medication) and item 4 (monitoring your medical conditions (eg, checking your blood pressure or blood sugar, monitoring your symptoms, etc)).

"Healthcare contacts and health information": item 5 (arranging appointments with health professionals), item 6 (seeing lots of different health professionals), item 7 (attending appointments with health professionals (eg, getting time off work, arranging transport, etc)) and item 8 (obtaining clear and up-to-date information about your condition).

"Coping ability": item 9 (making recommended lifestyle changes (eg, diet and exercise)) and item 10 (having to rely on help from family and friends).

Two items (9 and 10) had a uniqueness slightly larger than the recommended value of 0.6.

The partitioning into the three factors above was also found in a supplemental analysis, where the model for rotation was restricted to include maximum three factors (results not shown).

The three factors above were used as guide for the multifactor CFA.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix C

Three-factor CFA model

We used the same CFA model as in the main analyses with the exception that the items were now divided into three factors instead of one. The results are presented in table 4 and table 5.

Table 4. Factor loadings from confirmatory factor analyses with Satorra-Bentler adjustments

	One-factor model with four correlated error terms		**Three-factor model with four correlated error terms#**	
Please tell us how much difficulty you have with the following:	**Factor loadings**	**(95% CI)**	**Factor loadings**	**(95% CI)**
			Medication and self-monitoring	
1. Taking lots of medications	0.53	(0.50 to 0.55)	0.57	(0.55 to 0.60)
2. Remembering how and when to take medication	0.56	(0.53 to 0.59)	0.62	(0.59 to 0.65)
3. Collecting prescription medication	0.61	(0.58 to 0.64)	0.66	(0.63 to 0.68)
4. Monitoring your medical conditions (eg, checking your blood pressure or blood sugar, monitoring your symptoms, etc)	0.66	(0.63 to 0.68)	0.71	(0.69 to 0.74)
			Healthcare contacts and health information	
5. Arranging appointments with health professionals	0.75	(0.73 to 0.77)	0.78	(0.76 to 0.80)
6. Seeing lots of different health professionals	0.74	(0.73 to 0.76)	0.78	(0.76 to 0.80)
7. Attending appointments with health professionals (eg, getting time off work, arranging transport, etc)	0.67	(0.65 to 0.69)	0.69	(0.66 to 0.71)
8. Obtaining clear and up-to-date information about your condition	0.69	(0.68 to 0.71)	0.71	(0.69 to 0.73)
			Coping ability	
9. Making recommended lifestyle changes (eg, diet and exercise)	0.53	(0.51 to 0.55)	0.60	(0.57 to 0.62)
10. Having to rely on help from family and friends	0.57	(0.54 to 0.59)	0.63	(0.61 to 0.66)

Unweighted analyses.
One-factor model from the main analysis (for comparison) and supplementary three-factor model guided by the exploratory factor analysis.
#Correlations allowed between: i) questions 1 and 2; ii) questions 2 and 3; iii) questions 5 and 6; iv) questions 5 and 10.
Cov(Medication and self-monitoring ; Healthcare contacts and health information)=0.84.
Cov(Medication and self-monitoring ; Coping ability)=0.84.
Cov(Healthcare contacts and health information ; Coping ability)=0.82.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix C

Table 5. CFA models - global goodness-of-fit indices

	SBχ^2(df)	P value	SB-RMSEA	SB-CFI	SB-TLI	SRMR
One-factor, with four correlations between errors#	680.22(31)	<0.0001	0.041	0.952	0.930	0.036
Three-factor, with four correlations between errors#	345.27(28)	<0.0001	0.030	0.976	0.962	0.022

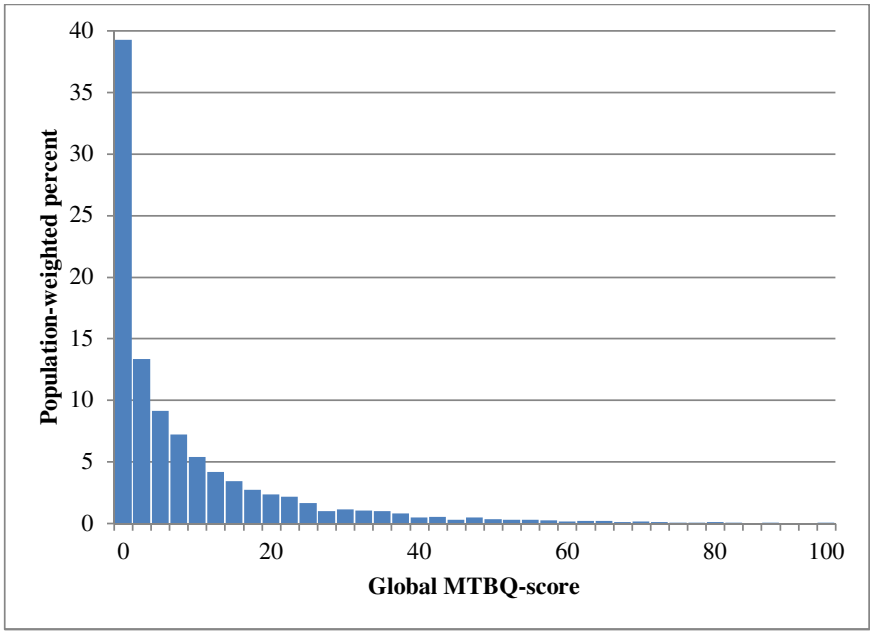
Unweighted analyses.
One-factor model from the main analysis (for comparison) and supplementary three-factor model guided by the exploratory factor analysis.
#Correlations allowed between: i) items 1 (taking lots of medications) and 2 (remembering how and when to take medication); ii) items 2 and 3 (collecting prescription medication); iii) items 5 (arranging appointments with health professionals), and 6 (seeing lots of different health professionals); iv) items 5 and 10 (having to rely on help from family and friends).
CFI, Comparative Fit Index; RMSEA, the root mean square error of approximation; SB, Satorra-Bentler adjusted; SRMR, standardised root mean square residual; TLI, Tucker-Lewis Index.

The overall coefficient of determination in the three-factor model was 0.94. In the one-factor model, it was 0.87.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix D

Appendix D: Population-weighted histogram of global MTBQ score, 2017 'How are you?' survey



Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix E

Appendix E: Factor loadings from confirmatory factor analyses with Satorra-Bentler adjustments

Please tell us how much difficulty you have with the following:	One-factor model with no correlated error terms		One-factor model with four correlated error terms#	
	Factor loadings	(95% CI)	Factor loadings	(95% CI)
1. Taking lots of medications	0.54	(0.52 to 0.57)	0.53	(0.50 to 0.55)
2. Remembering how and when to take medication	0.59	(0.56 to 0.62)	0.56	(0.53 to 0.59)
3. Collecting prescription medication	0.62	(0.59 to 0.64)	0.61	(0.58 to 0.64)
4. Monitoring your medical conditions (eg, checking your blood pressure or blood sugar, monitoring your symptoms, etc)	0.66	(0.63 to 0.68)	0.66	(0.63 to 0.68)
5. Arranging appointments with health professionals	0.76	(0.74 to 0.78)	0.75	(0.73 to 0.77)
6. Seeing lots of different health professionals	0.77	(0.75 to 0.79)	0.74	(0.73 to 0.76)
7. Attending appointments with health professionals (eg, getting time off work, arranging transport, etc)	0.66	(0.64 to 0.69)	0.67	(0.65 to 0.69)
8. Obtaining clear and up-to-date information about your condition	0.69	(0.67 to 0.71)	0.69	(0.68 to 0.71)
9. Making recommended lifestyle changes (eg, diet and exercise)	0.52	(0.50 to 0.54)	0.53	(0.51 to 0.55)
10. Having to rely on help from family and friends	0.53	(0.51 to 0.56)	0.57	(0.54 to 0.59)

Unweighted analyses.
#Correlations allowed between: i) questions 1 and 2; ii) questions 2 and 3; iii) questions 5 and 6; iv) questions 5 and 10.

Danish validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ) and findings from a population health survey: a mixed-methods study

Appendix F

Appendix F: Relaxation of the proportional odds assumption for those variables that violated the assumption

	Unadjusted OR (95% CI)*			**Adjusted OR (95% CI)*†**		
	Model A: No burden vs low+ medium+high burden	Model B: No+low burden vs medium+high burden	Model C: No+low+medium burden vs high burden	Model A: No burden vs low+ medium+high burden	Model B: No+low burden vs medium+high burden	Model C: No+low+medium burden vs high burden
Sociodemographic factors						
Age, years (ref.=25-34)						
55-64				**0.34 (0.28 to 0.42)**	**0.33 (0.27 to 0.41)**	***0.25 (0.20 to 0.33)***
Country of origin (ref.=Denmark)						
Other	**1.32 (1.08 to 1.61)**	***1.77 (1.46 to 2.15)***	***2.15 (1.70 to 2.71)***	**0.75 (0.59 to 0.96)**	*1.12 (0.88 to 1.43)*	*1.27 (0.93 to 1.76)*
Educational level (ref.=Medium (11-14 years))						
High (15+ years)	**1.19 (1.08 to 1.31)**	*0.99 (0.89 to 1.11)*	***0.83 (0.72 to 0.97)***	**1.31 (1.17 to 1.47)**	*1.13 (1.00 to 1.27)*	*0.93 (0.78 to 1.12)*
Employment status (ref.=Employed or student)						
Permanently out of work	**0.71 (0.65 to 0.77)**	***0.82 (0.74 to 0.90)***	*0.99 (0.86 to 1.12)*			
Self-reported long-term conditions						
Stroke	**1.85 (1.46 to 2.35)**	***2.48 (2.00 to 3.06)***	***3.18 (2.50 to 4.04)***			
Scores of health measures‡						
Mean self-rated health score (SD)§	**0.38 (0.36 to 0.40)**	***0.33 (0.31 to 0.35)***	***0.28 (0.26 to 0.31)***	**0.41 (0.39 to 0.44)**	***0.37 (0.35 to 0.40)***	***0.33 (0.30 to 0.37)***
Mean PCS score (SD)¶	**0.92 (0.92 to 0.93)**	***0.92 (0.91 to 0.92)***	***0.91 (0.90 to 0.91)***			
Mean MCS score (SD)¶	**0.91 (0.90 to 0.91)**	***0.90 (0.89 to 0.90)***	***0.89 (0.89 to 0.90)***	**0.92 (0.91 to 0.92)**	***0.91 (0.90 to 0.91)***	***0.90 (0.90 to 0.91)***
Mean PSS score (SD)	**1.12 (1.12 to 1.13)**	***1.16 (1.15 to 1.16)***	***1.18 (1.16 to 1.19)***	**1.10 (1.09 to 1.11)**	***1.13 (1.12 to 1.14)***	***1.14 (1.13 to 1.16)***

Bold values indicate that the covariate is associated with a statistical significant OR in the respective model (model A, B or C) at the 5% level. The ORs are presented with their respective CIs.

Italic values indicate that the OR associated with a specific covariate is statistically significant different between the respective model (model B or C) compared to the OR from model A at the 1% level (P values for this test is not shown for simplicity).

*Partial proportional odds models weighted to represent the population of the Central Denmark Region, aged 25+ years, in treatment.

†Adjusted for age, sex, country of origin, educational attainment, employment status, marital status, living with child(ren) aged 0-15 years, deprivation level, individual long-term conditions and multimorbidity.

‡The scores of health measures are included in the regression analyses as continuous variables. Hence, the odds ratios represent the odds of a higher level of treatment burden category versus the current or lower level treatment burden category for a one unit increase in the health measure score.

§Single question: 'In general, would you say your health is: excellent (5), very good (4), good (3), fair (2), poor (1)'?

¶Based on the SF-12v2-questionnaire.

COPD, chronic obstructive pulmonary disease; MCS, Mental Component Summary; PCS, Physical Component Summary; PSS, Perceived Stress Scale.