Prediction model study focusing on eHealth in the management of urinary incontinence: the Personalised Advantage Index as a decision-making aid

Anne Martina Maria Loohuis, Huibert Burger, Nienke Wessels, Janny Dekker, Alec GGA Malmberg, Marjolein Y Berger, Marco H Blanker, Henk van der Worp

ABSTRACT

Objective To develop a prediction model and illustrate the practical potential of personalisation of treatment decisions between app-based treatment and care as usual for urinary incontinence (UI).

Design A prediction model study using data from a pragmatic, randomised controlled, non-inferiority trial.

Setting Dutch primary care from 2015, with social media included from 2017. Enrolment ended on July 2018.

Participants Adult women were eligible if they had ≥2 episodes of UI per week, access to mobile apps and wanted treatment. Of the 350 screened women, 262 were eligible and randomised to app-based treatment or care as usual; 195 (74%) attended follow-up.

Predictors Literature review and expert opinion identified 13 candidate predictors, categorised into two groups: Prognostic factors (independent of treatment type), such as UI severity, menopausal state, vaginal birth, general physical health status, pelvic floor muscle function and body mass index; and modifiers (dependent on treatment type), such as age, UI type and duration, impact on quality of life, previous physical therapy, recruitment method and educational level.

Main outcome measure Primary outcome was symptom severity after a 4-month follow-up period, measured by the International Consultation on Incontinence Questionnaire the Urinary Incontinence Short Form. Prognostic factors and modifiers were combined into a final prediction model. For each participant, we then predicted treatment outcomes and calculated a Personalised Advantage Index (PAI).

Results Baseline UI severity (prognostic) and age, educational level and impact on quality of life (modifiers) independently affected treatment effect of eHealth. The mean PAI was 0.99±0.79 points, being of clinical relevance in 21% of individuals. Applying the PAI also significantly improved treatment outcomes at the group level.

Conclusions Personalising treatment choice can support treatment decision making between eHealth and care as usual through the practical application of prediction modelling. Concerning eHealth for UI, this could facilitate the choice between app-based treatment and care as usual.

Trial registration number NL4948t.

INTRODUCTION

Randomised controlled trials (RCTs) provide evidence of treatment effects at a group level, but they fail to provide the individual-level predictive information needed to optimise treatment in a given patient. This is especially relevant when two treatments show only marginal differences in effect at the group level, as occurs in a non-inferiority design, where the added value of personalised treatment decision might be greater. A prediction model for treatment outcome,
based on a patient’s individual characteristics, may facilitate the personalisation of treatment decisions. Different approaches to the development of clinical decision support tools informed by prediction models have been published in epidemiological and statistical literature, being developed for various disorders. In mental healthcare, where treatment options for depression often show comparable effectiveness and marked individual variability, the Personalised Advantage Index (PAI) has shown utility. This method predicts individualised outcomes at follow-up, where treatment options for depression often show non-inferiority at a group level, but we equally found that individual outcomes at follow-up varied from ‘much worse’ to ‘very much better’ in both treatment groups. Previously, higher age, treatment expectations and disease severity were reported to predict better outcomes for UI when using eHealth. Although a given patient and caregiver could weigh these separate characteristics when making a treatment decision, it would be much more informative to know what specific outcomes one can expect from the available treatment options. We are unaware of the PAI having been applied to treatment decisions concerning eHealth.

In this study, we used existing RCT data to develop a prediction model and illustrate how the personalisation of treatment decisions affects the choice between app-based treatment and care as usual. We also studied the practical potential of this approach in women with stress, urgency or mixed UI. First, we built a prediction model for the outcomes of app-based treatment and care as usual for UI, and we used this to predict outcomes given the actual treatment received (factual) and the hypothetical outcome of the treatment that was not received (counterfactual). Second, we used the PAI to identify the optimal treatment and to quantify its added benefit in individual participants. Third, we assessed the clinical relevance of any benefit and whether using the PAI improved treatment outcomes at the group level.

METHODS

Data source and study design

We used data from the UrinControl-trial, a pragmatic, non-inferiority RCT of women with stress, urgency or mixed UI who received either an app-based treatment or care as usual via their general practitioner (GP). The trial design, the development and content of the app, and the clinical results have been published previously. The original trial reported the non-inferiority of app-based treatment to care as usual at a group level. Baseline characteristics and outcome measures were based on data collected through validated questionnaires and a physical examination by a GP trainee. In this study, we use these data to build a prediction model, predict treatment outcomes at an individual level and calculate the PAI.

Participants

Participant enrollment took place from July 2015 to July 2018, with follow-up ending on 20 December 2018. We recruited participants in the north of the Netherlands via 88 GPs from 31 practices, and through social media and the lay press. Adult women were eligible if they had ≥2 episodes of self-reported stress, urgency or mixed UI per week, a wish to be treated, and access to a smartphone or tablet. The exclusion criteria were as follows: urinary tract infection, overflow or continuous UI, indwelling urinary catheter, urogenital malignancy, pregnancy or recent childbirth (<6 months ago), treatment for UI in the previous year, previous surgery for UI, terminal or serious illness, and cognitive impairment, psychiatric illness or the inability to complete a questionnaire in Dutch. The present analyses used the pretreatment data and the outcome data at 4 months for all women included in the original study.

Treatments

App-based treatment consisted of a step-by-step programme for the self-management of UI, with content based on relevant Dutch GP and international guidelines. Care as usual comprised referral to a continence nurse, a pelvic physical therapist; pelvic floor muscle training and/or bladder training; prescribing of a pessary, drugs or absorbent products; and referral to a continence nurse, a pelvic physical therapist or secondary care.

Outcome

The outcome predicted by the model was UI severity after 4 months of treatment, which we labelled the end-Urinary Incontinence Short Form (UISF) score. This continuous score was measured by the International Consultation on Incontinence Questionnaire, Urinary Incontinence Short Form (ICIQ-UISF), a questionnaire measuring the self-reported frequency, severity and impact on daily life of UI. Scores ranged from 0 to 21, with higher scores indicating worse incontinence. Data analysts were blinded to the treatment arm at the time of analysis.

Predictors

We identified candidate predictors based on literature search and expert opinion. PubMed was searched for predictors of conservative UI treatment and eHealth

We calculated the maximum number of parameters needed to build the model according to the guidance of Riley et al, based on a clinical prediction model with a continuous outcome and a known sample size of 262 participants. Given a mean 9.9±3.3-point UI severity score from our trial population and an anticipated $R^2$ of 0.6 from Nystroem et al, we calculated that a maximum of 28 parameters were needed to build the model.

Data were missing for the outcome measure and one predictor, which we accommodated by multiple imputation under the assumption of data being missing at random. We assessed the missing data mechanism by looking at patterns and predictors of missingness to substantiate assumptions of being missing at random or not at random. All variables that predicted missingness of a certain variable were included in the imputation model together with all variables from the analyses.

All statistical analysis was performed using IBM SPSS for Windows, V.26.0 (IBM) and R. We performed multiple imputation in R, using the MICE package and constructed 50 imputed datasets.

Development and validation

We developed a model to predict the treatment outcome based on prognostic factors and modifiers, assessed optimism by internal validation and applied this as a correction. This model was used to predict two outcomes for each participant: (1) for the actual treatment the patient received and (2) for the counterfactual treatment to which the patient was not allocated. We used these to construct and calculate the PAI, before assessing its benefits at individual and group levels.

Step 1: development of the prediction model

We investigated multicollinearity in the non-imputed dataset by correlation matrix, which revealed that none of the candidate predictors correlated highly with another ($r>0.8$). As described by Kraemer et al, continuous predictors were centred by subtracting the median and dichotomous variables were set at 0.5 and −0.5. The end-USIF score was predicted by linear regression, with the potential prognostic factors entered as main effects and the potential modifiers entered as both main effects and terms representing their interactions with treatment. We used a stepwise, backward elimination strategy, excluding variables from the model based on an alpha of 0.25. Predictors selected in at least 50% of the imputed datasets were included in the final model. We forced treatment type and the main effects of every included interaction into the final model irrespective of their significance. Model performance was assessed by $R^2$, goodness-of-fit and calibration slope. The 95% CIs are reported as appropriate.

Step 2: internal validation of the prediction model

Stability of the regression coefficients, inclusion percentages and the mean adjusted $R^2$ was assessed across 500
bootstrapped samples. We examined precision with the true error (mean observed score minus mean predicted score) and the SE. To correct for overoptimism, we equally applied uniform shrinkage to the final model coefficients.32

Step 3: construction of the PAI
Having determined the predictors of differential response, a model can be constructed to generate treatment recommendations, making use of the PAI.4 Given our aim to study the practical potential of this index, we focused on clinical utility over technical detail (figure 1).4 5

Prediction of individual outcomes
For each patient, we predicted the end-UISF score for app-based treatment and for care as usual by completing the model twice with the patient’s observed values: once with the value of app-based treatment (−0.5) and once with the value of care as usual (0.5). This predicted the end-UISF score for the treatments the patient received (factual score) and did not receive (counterfactual score).

To predict individual end scores, we split the sample into five equal groups and used a linear regression model with sampling weights based on data from four groups to predict end scores in the targeted group. This fivefold cross-validation reduced the risk of overfitting by avoiding the inclusion of an individual’s own data when estimating the relevant regression coefficients.

Interpretation of individual outcomes
Three end-UISF scores were documented for each patient: (1) the observed score after receiving the randomised treatment in the trial, (2) the predicted score for app-based treatment and (3) the predicted score for care as usual. The lowest of the two predicted scores was the optimal treatment (figure 1).

Step 4: assessment of the PAI
Assessment of individual benefit
For each patient, we then calculated the PAI as a measure of the benefit of one treatment over the other and assessed its magnitude (ie, clinical relevance). The PAI was the difference between the highest and lowest predicted score. Based on a difference of 1.58 points having previously been defined as the minimum clinically important difference for the ICIQ-UISF,25 optimal treatment with a PAI higher than this was expected to have a noticeable benefit for the patient.

Assessment of improvement at the group level
Finally, we assessed whether treatment personalisation using the PAI significantly and substantially improved

<p>| Table 1  Baseline characteristics of all participants with urinary incontinence |
|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (n=262)</th>
<th>App treatment (n=131)</th>
<th>Care as usual (n=131)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prognostic factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severity UI at baseline*</td>
<td>9.9±3.3</td>
<td>9.5±3.2</td>
<td>10.3±3.4</td>
</tr>
<tr>
<td>Body mass index (kg/m²)*</td>
<td>27.8±5.3</td>
<td>27.6±5.5</td>
<td>28.0±5.2</td>
</tr>
<tr>
<td>Postmenopausal status, yes</td>
<td>123 (47.1%)</td>
<td>64 (49.2%)</td>
<td>59 (45.0%)</td>
</tr>
<tr>
<td>Vaginal births,≥1</td>
<td>216 (82.8%)</td>
<td>111 (85.4%)</td>
<td>105 (80.2%)</td>
</tr>
<tr>
<td>Pelvic floor muscle function</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal activity</td>
<td>84 (32.1%)</td>
<td>44 (33.6%)</td>
<td>40 (30.5%)</td>
</tr>
<tr>
<td>Overactive</td>
<td>44 (16.8%)</td>
<td>18 (13.7%)</td>
<td>26 (19.8%)</td>
</tr>
<tr>
<td>Underactive</td>
<td>134 (50.8%)</td>
<td>69 (52.7%)</td>
<td>65 (49.6%)</td>
</tr>
<tr>
<td>General physical health status*</td>
<td>74±20</td>
<td>73±20</td>
<td>75±21</td>
</tr>
<tr>
<td>Modifiers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, (years)</td>
<td>52.2±11.6</td>
<td>53.2±12.8</td>
<td>51.3±10.3</td>
</tr>
<tr>
<td>Educational level, higher</td>
<td>107 (52.7%)</td>
<td>58 (54.2%)</td>
<td>49 (51.0%)</td>
</tr>
<tr>
<td>Duration of UI (years)*</td>
<td>7 (4–14)</td>
<td>7 (4–15)</td>
<td>8 (4–13)</td>
</tr>
<tr>
<td>UI impact on quality of life*</td>
<td>33.6±8.0</td>
<td>33.9±8.3</td>
<td>33.4±7.8</td>
</tr>
<tr>
<td>Type of UI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress</td>
<td>180 (68.7%)</td>
<td>87 (66.4%)</td>
<td>93 (71.0%)</td>
</tr>
<tr>
<td>Urgency</td>
<td>82 (31.3%)</td>
<td>44 (33.6%)</td>
<td>38 (29.0%)</td>
</tr>
<tr>
<td>Previous physical therapy for UI, yes</td>
<td>66 (25.3%)</td>
<td>31 (23.8%)</td>
<td>35 (26.7%)</td>
</tr>
<tr>
<td>Recruitment type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General practitioner</td>
<td>152 (58%)</td>
<td>76 (58.0%)</td>
<td>76 (58.0%)</td>
</tr>
<tr>
<td>Lay press or social media</td>
<td>110 (42%)</td>
<td>55 (42.0%)</td>
<td>55 (42.0%)</td>
</tr>
</tbody>
</table>

Prognostic factors predict outcomes irrespective of treatment type. Modifiers predict outcomes dependent on the treatment (modifier). Values are presented as means±SD deviation, percentages or medians (IQR).

*N was lower: missing data of one baseline assessment, three baseline questionnaires and educational level were assessed at follow-up. UI, urinary incontinence.
treatment outcomes at a group level (ie, the usefulness of the PAI as a tool to improve treatment selection and thereby effectiveness). Using the observed outcome scores from the trial, we compared patients who randomly received an optimal treatment with those who randomly received a non-optimal treatment. Randomisation for this comparison allowed causal interpretation at the group level because we built the model on a separate selection of participants (using fivefold cross-validation) and because the predicted end-UISF scores were not tied to the randomisation or treatment received.

Patient and public involvement
We involved patients, the public and professionals from the start of the study.10 Each group provided feedback on the study design, assessed the app in the development phase, provided feedback during the trial phase and have been informed of previous results by email. They have been informed of previous results by email and will be informed of the current results after publication. To facilitate the process of informing the public, we have produced plain-language summaries in text, illustration and video formats for dissemination on social media and our website (www.urincontrol.nl).

RESULTS
Participants
We included data for 262 women who participated in the trial (table 1). The only remarkable difference was a lower severity of UI in the app-based treatment group. At 4 months, 195 women (74.4%) had reported end-UISF scores, resulting in 67 cases of missing data for the end-UISF score and educational level. The outcome variable was missing at random and predicted by a younger age, a higher body mass index and no prior treatment, but not by severity of incontinence.7

Development of the prediction model
Table 2 shows the variables included in the final model. The model explained 46% of the variance in predicting the outcome measure (R² 0.46; 95% CI 0.36 to 0.55)). The mean difference between observed and predicted outcomes in the original data—that is, the goodness-of-fit—was 0.015 (95% CI −0.308 to 0.278), showing a calibration intercept at −0.06 and a calibration slope of 1.01 (online supplemental figure 1). A lower end-UISF score, indicating a better treatment outcome, was predicted by care as usual, lower baseline severity of UI, and lower impact of UI on quality of life. Success of app-based treatment was associated with higher age, higher impact of UI on quality of life and higher educational level. Success of care as usual was associated with lower educational level.

Internal validation
Regression coefficients and inclusion percentages across the bootstrapped samples were stable. The mean R² (% explained variance) was 0.455 (95% CI 0.357 to 0.547) after bootstrapping (online supplemental table 3). The uniform shrinkage factor calculated by bootstrapping was small with a factor of 0.98 (table 2). The true error was 1.85 (values plotted in online supplemental figure 2) and the SE was 0.15.

Personalised Advantage Index
At the group level, the mean change in the unimputed observed UISF score after 4 months indicated a symptoms improvement of −2.35±3.05 points. The change of symptom score varied from −15 to 6 points among patients.

Individual observed and predicted outcome scores
The observed scores showed a mean end-UISF of 7.58±3.46 points (range 0–18). The mean predicted optimal and non-optimal scores per patient were 7.15±2.46 points
Individual PAl scores and their clinical relevance legend: the figure shows the individual variability of treatment response above and below the minimum clinical important difference of 1.58. PAl, Personalised Advantage Index; MCID, minimum clinically important difference.

(range 0–13) and 8.14±2.52 points (range 2–16), respectively.

Individual benefit
The PAl showed a mean benefit of 0.99±0.79 points for the optimal treatment over the non-optimal treatment, which ranged from 0.02 to 4.21 points at the individual level (figure 2). This difference was clinically relevant at ≥1.58 points for 55 patients (21%). Online supplemental table 4 shows a comparison of baseline characteristics between 55 patients with a clinically relevant PAl and the other patients.

Improvement on group level
Finally, we compared the observed trial outcomes of patients receiving optimal (n=135; 51%) and non-optimal (n=127; 49%) treatments, which had mean scores of 7.01±3.33 points and 8.20±3.51 points, respectively. The observed difference in means between the randomised optimal and non-optimal treatment was statistically significant with a mean difference of 1.19 points (95% CI 0.355 to 2.021).

DISCUSSION
Statement of principal findings
We illustrated a method for predicting optimal treatment and for quantifying its benefit compared with non-optimal treatment at the individual patient level when eHealth is being considered for UI management. Four baseline characteristics, namely UI severity (a prognostic factor) and age, educational level and impact of UI on quality of life (modifiers), were identified as suitable for helping with decisions in this model. The mean advantage according to the PAl was 0.99 points, and it exceeded the threshold for clinical relevance of 1.58 in 21% of individuals. Applying the PAl to facilitate decision making also significantly improved treatment outcomes at the group level, which may be relevant when considering other measures of quality with this treatment, such as cost-effectiveness. To our knowledge, this is the first study to have translated established methods for predicting treatment outcomes from mental health and somatic disease settings to an eHealth setting.

Strengths and weaknesses of the study
We developed a model for predicting the treatment option most likely to improve UI symptoms in individuals and assessed the clinical relevance of that prediction. The use of data from a pragmatic RCT with a representative first-line population make that data well suited to developing a prediction model for personalising treatment decisions. The model is also useful because the predictors are both easily reproduced (age and answers to validated questions) and readily available in clinical practice. Other strengths are the use of a patient-centred outcome measure, the selection of predictors based on literature and expert opinion, the inclusion of both prognostic factors (treatment independent) and modifiers (treatment dependent), the power of the prediction study and the minimal overfitting of the model (shrinkage factor=0.98).

There are several important limitations that should also be considered. First, the prediction model required external validation via a sample comparing app-based treatment to care as usual; however, such a sample does not exist for UI. Internal validation only confirmed the stability of development and performance of the model. Second, the explained variance of 46% was moderate, being similar to that reported for other eHealth models for UI (range 30%–61.4%). Third, the true error was 1.85 points in our sample, which is larger than the mean PAl (0.99 points) and threshold for clinical relevance (1.58 points), possibly indicating low precision for personalised predictions and probably affecting the estimation of the magnitude of an individual’s advantage. Performance and precision could be increased by adding stronger predictors that interact with eHealth treatment to the model. Despite a thorough search for candidate predictors, those interacting with eHealth treatment could have been missed, especially given that literature on this topic is scarce. Finally, we could not include some variables identified by literature search and expert opinion, such as the eHealth literacy and treatment expectations of participants, because these were missing from our dataset.

Strengths and weaknesses in relation to other studies and key differences
Our approach of predicting the best treatment option for an individual shows important features and challenges of predictive heterogeneity of treatment effect analysis. Our methods fall under the larger ‘effect modelling’ approach (as opposed to risk modelling), meaning that a term for treatment assignment and interactions between treatment and baseline covariates are included in the prediction model, as described in a recent State of the art review. Disaggregation of the overall results in our
pragmatic trial appeared to improve treatment effects on the individual and population level. However, we also encountered barriers linked to effect modelling described in the review; for example, a priori modifiers that were not yet well established, limited statistical power, multiple testing problems and the evaluation whether a particular prediction-decision strategy (in our trial for example clinical relevance threshold) would optimise the net benefit in the general population. The PAI predicted clinically relevant improvement in 60% of patients with depression for the choice between antidepressant medication and cognitive behavioural therapy. Compared with the mental health setting, clinically relevant improvement was only predicted in 21% of our cohort, possibly because there is less existing knowledge or less identifiable variation in treatment effect for UI. To date, however, we are unaware of any other studies having assessed and compared the interaction of predictors for an eHealth treatment and care as usual. We believe this type of assessment is essential to strengthen treatment-specific outcome predictions and to optimise clinical decision making for personalised medicine. Lindh et al, for example, showed that higher age predicted greater treatment success with eHealth for UI, but this only considered their total sample (internet-based treatment and controls) and did not assess treatment interactions. In our study, the interaction of age with treatment type implies that a higher age may favour app-based treatment over care as usual. This new information is relevant to both researchers and clinicians because it runs counter the general expectation that eHealth is better suited to younger patients.

Possible mechanisms and explanations for findings
The predictors in our model should not be interpreted as strict causal or etiological factors for UI symptoms. The present data analysis was designed specifically to identify a set of variables that had high predictive accuracy in combination, rather than to unravel the causal factors influencing UI symptom severity at follow-up. However, the predictors that remained in the model had high a priori predictive value and are plausible causal factors.

In the developed model, increased age, educational level and impact of UI on quality of life predicted a better treatment outcome for app-based treatment compared with care as usual. Educational level had the greatest modifying effect, with a higher level associated with benefit from app-based treatment and a lower level associated with care as usual. This is likely to reflect differences in health, eHealth literacy and self-efficacy, but it could also reflect the app’s design (eg, there are lengthy sections of text or instructions that may be too difficult to understand) or better adaptability by a given healthcare professional to a patient’s need for support.

Other studies indicate that lower health literacy is associated with poorer health outcomes and with difficulties using eHealth effectively. A mobile app has the potential to be tailored to specific users, such as those with low literacy and may bridge this gap. Given that the content of our app was not tailored to users with low literacy, we will develop it further to improve its availability, readability and usability. Furthermore, we plan to add improved technological and practical support, specifically targeting users with low literacy.

Potential implications for clinicians or policy-makers
Prediction modelling at a group level only allows patients and caregivers to guess how a given characteristic influences treatment outcomes at an individual level. The PAI helps to correct this by quantifying the expected outcomes and benefits of an optimal treatment over its alternative given an individual’s characteristics. The results are easy to interpret and can inform decisions immediately.

Our model requires further development and validation, but in the meantime, we believe it can be used in clinical practice. Indeed, using the tool is certainly superior to the current situation where no support is available and where its use will pose little risk to the patient if the prediction is wrong (ie, the options are non-inferior, but its use could improve outcomes). The PAI could also be implemented in clinical practice with ease, either within the app itself or on a patient information website, where the necessary prognostic factors and modifiers can be entered by users to predict the option most likely to be of benefit. This approach could be especially helpful for shared decision making and could be used to guide patients who wish to consider using a freely accessible eHealth intervention when they experience barriers seeking help from a caregiver. Currently, these patients often start to use an available app with no knowledge of what to expect.

Unanswered questions and future research
We missed important predictors by not anticipating the present analysis at the inception of our trial. Therefore, we recommend that eHealth researchers consider adding a method for personalising treatment decisions to allow them to consider and include all relevant predictors. This is especially relevant if researchers are conducting a (pragmatic) RCT, which otherwise provides the perfect foundation for this method. If more eHealth researchers conducted similar research, we may see a large-scale improvement in clinical decision making, treatment outcomes and our knowledge of the predictors that interact with eHealth treatment.

External validation of the model in the present study is needed, but this is complicated by the lack of a suitable sample. More validation samples may be available for researchers applying this method to other eHealth settings where there is a greater body of research comparing eHealth to care as usual (eg, obesity and diabetes). Finally, an impact study comparing treatment outcomes for groups with and without this decision support tool would be of interest.
CONCLUSION

Prediction modelling can directly support decisions to personalise treatment when choosing between eHealth and care as usual. We applied this principle to an eHealth treatment for UI and, despite our model having only moderate predictive performance and still requiring external validation, we demonstrated its practical potential.

Twitter Anne Martina Maria Loohuis @loohuisanne and Marco H Blanker @Marco_Blanker

Acknowledgements We thank the participating general practices for their ongoing support, as well as all the participants for their invaluable contributions to this study. Special thanks is due to patients involved in the development of the app and to those at the Bekkentodemonfall patient organisation. Finally, we thank Dr Robert Sykes (www.doctored.org.uk) for providing editorial services.

Contributors AMML collected the data, did the analysis and wrote the paper; HB designed the study and contributed to the analysis and writing of the paper; NW collected the data and contributed to the writing of the paper; JD designed the study, acquired the funding, and contributed to the writing of the paper; AGGAM assisted in the study design, the content of the app and contributed to the writing; MYB assisted in the study design and contributed to the writing of the paper; MHB designed the study, acquired the funding, was project leader, contributed to the analysis and contributed to the writing of the paper; HvdW contributed to the analysis and contributed to the writing of the paper; and MHB is guarantor. The corresponding author attests that all listed authors meet the criteria for authorship and that no others meeting the criteria have been omitted.

Funding This work was supported by a grant from ZonMw, The Dutch Organisation for Health Research and Development (project number: 837001508) and subfunded by a grant from the P.W. Boer foundation. The study won the Professor Huygen award 2016 for best study proposal in general practice, which included additional funding.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Consent obtained directly from patient(s)

Ethics approval This study involves human participants and was approved by the Medical Ethical Review board of the University Medical Center Groningen (Netherlands) (METHC-number: 2014/574) approved this study. All participants gave written informed consent. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

Individual participant data that underlie the results reported in this article, after the identification, will be available including data dictionaries. Data are available to investigators who provide a methodologically sound proposal for analyses to achieve aims in the approved proposal. The available period begins 9 months after publication and ends 36 months following article publication and after approval of a proposal. Additional information available is the study protocol. There are no additional restrictions on the use of the data. The data are available from MHB, m.h. blanken@umcg.nl.

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 peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

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

ORCID iDs
Anne Martina Maria Loohuis http://orcid.org/0000-0001-6582-1502
Huibert Burger http://orcid.org/0000-0001-6484-0349
Marco H Blanker http://orcid.org/0000-0002-1086-8730
Henk van der Worp http://orcid.org/0000-0001-5545-4155

REFERENCES


