AARHUS UNIVERSITY

Patient-reported outcome measures in remote outpatient follow-up

PhD dissertation

Liv Marit Valen Schougaard

Health
Aarhus University
2020

Patient-reported outcome measures in remote outpatient follow-up

PhD dissertation

Liv Marit Valen Schougaard

Health

Aarhus University

Unit for Patient-reported Outcomes, WestChronic/AmbuFlex,

Regional Hospital West Jutland

"I consider AmbuFlex to be a useful tool
that enables us to offer the patients
an individual flexible contact,
adjustable to the patients' needs."

Per Sidenius MD,
Head of Department of Neurology,
Aarhus University Hospital
2009–2017

Supervisors

Professor Niels Henrik Hjøllund, MD, PhD
WestChronic, Occupational Medicine, University Research Clinic, Aarhus University, Denmark
Department of Clinical Epidemiology, Aarhus University Hospital, Denmark

Associate Professor Annette de Thurah, MPH, PhD
Department of Rheumatology, Aarhus University Hospital, Denmark
Department of Clinical Medicine, Aarhus University, Denmark

Associate Professor Jakob Christensen, MD, PhD
Department of Neurology, Aarhus University Hospital, Denmark
Department of Clinical Medicine, Aarhus University, Denmark

Professor Kirsten Lomborg, MSc Nursing, PhD
Department of Clinical Medicine, Aarhus University, Denmark

Evaluation Committee

Associate Professor Palle Bager, PhD (Chairman)
Department of Hepatology and Gastroenterology, Aarhus University Hospital, Denmark

Professor of Medicine Claire Snyder, PhD
Division of General Internal Medicine, Johns Hopkins School of Medicine, USA

Professor Marjolein Memelink Iversen, PhD
Department of Health and Caring Sciences, Western Norway University of Applied Sciences, Norway

PREFACE AND ACKNOWLEDGEMENTS

This PhD dissertation is based on one of the first AmbuFlex solutions developed for use in outpatients with epilepsy. In 2007, Niels Henrik Hjøllund started developing a generic web-system called AmbuFlex which aimed to use patient-reported outcome (PRO) measures in clinical practice. In 2011, a highly engaged leadership including Per Sidenius and Mette Abildgaard Høi from the Department of Neurology at Aarhus University Hospital decided to use the AmbuFlex system in outpatients with epilepsy. I was employed as a project coordinator and was privileged to cooperate with a highly engaged clinical team from the department. The work done by this group of people has been the foundation for the upcoming AmbuFlex solutions. In 2013, it was decided to implement three new AmbuFlex solutions yearly in Central Denmark Region. Following, in 2017, a joint steering group across all five Danish Regions was established, and as of January 2020, AmbuFlex has been implemented in 35 different patient groups in Danish hospitals.

The need for research in this field was raised from the very start and one of our research collaborators, Erik Riiskjær, made the quite incisive remark: *"It seems to me that practice has overtaken research."* In 2014, the ideas underlying this PhD dissertation were conceived. I was privileged to have the opportunity to design the project together with my main supervisor Niels Henrik Hjøllund. Per Sidenius, Annette de Thurah, and Kirsten Lomborg quickly came along and have contributed profoundly to the work presented in this dissertation. Along with my PhD project, Kirsten Lomborg was the main supervisor for a qualitative PhD project that was defended in 2018 by Caroline Mejdahl.

This PhD dissertation is now a final chapter in a nearly 9-year long journey I started when I was employed as a project coordinator at AmbuFlex at the Regional Hospital of West Jutland in 2011. At times the journey has been challenging and time-consuming, but it is with great pleasure that I look back at all the knowledge and skills I have gained through this journey and the PhD education. Fortunately, I do not feel this to be a final chapter, but rather the beginning of a new one.

In this period, I have had the privilege of working with a number of talented and inspiring people, and many deserve to be mentioned for their contribution to this project.

First and foremost, I would like to thank patients, clinicians, and the management at the three Departments of Neurology in Central Denmark Region. I owe a very special thanks to Lene Elbek Overbeck, Anette Møller Clausen, Birgitte Jensen Hanner, Mette Abildgaard Høi, and Per Sidenius for being PRO champions in a busy clinical practice. Apart from being PRO champions, they have also supported my PhD project, and they have never closed the door to my endless stream of enquiries.

I greatly acknowledge the financial support provided by the Central Denmark Regions Health Research Foundation, Aarhus University, and TrygFonden.

I would like to thank my great team of supervisors. Niels Henrik Hjøllund, my main supervisor, for introducing me to the world of patient-reported outcomes. Your

knowledge in this field has been inspiring, and I have learned a lot from the many discussions we have had. Your feedback has been pivotal throughout the entire process. I am indebted to Annette de Thurah, my co-supervisor and mentor, for always having an open door, and for encouraging, supportive, and challenging talks and to Kirsten Lomborg, my other co-supervisor, for her vast knowledge and constructive feedback. Unfortunately, Per Sidenius, who had a great part in the initial design of the PhD project, passed away all too soon. Jakob Christensen agreed to take up the work of supervision and has been a great help with his advice related to epilepsy and methodology.

I also owe a very special thanks to a handful of other researchers who have contributed to this PhD project with their knowledge: David Høyrup Christiansen for sharing his expertise in psychometric properties, Helle Terkildsen Maindal for introducing me to the concept of health literacy, Trine Nøhr Winding for her vast experience with Statistics Denmark, and Karin Biering for sharing her expertise in missing data and multiple imputation. Moreover, I would like to thank Jesper Medom Vestergaard for the greatest data management support and Morten Frydenberg for excellent statistical guidance. I also would like to thank all the current and previous fellow PhD students at the Department of Occupational Medicine in Herning. I have benefitted greatly from the supportive and humorous environment at Nord.

Thanks to Tone Merete Norekvål and colleagues for hosting me for an inspiring stay at the Centre on Patient-reported Outcome Data at Helse Bergen and Haukeland University Hospital in Norway. I am grateful for all the inspiring discussions, and I look forward to future collaboration.

I wish to thank all my great colleagues at AmbuFlex for their ongoing interest in my PhD project. I owe a special thanks to Louise Pape Larsen for supportive leadership, Kasper Bjerg for ongoing technical support, Anne Jessen for excellent feedback, and Maiken Norup Petersen for help with layout of figures. In particular, thanks to fellow PhD student Birgith Engelst Grove for continuous support and lots of fun in our office. I owe a special thanks to Caroline Mejdahl, my friend and previous PhD student fellow and colleague at AmbuFlex. It has been a privilege sharing ups and downs with you—thanks for all the wonderful conversations during the entire process.

Finally, I would like to thank my friends and family. Thanks for lots of fun and reminding me about the most important things in life. A warm thanks goes especially to my parents Frøydis and Oddbjørn Valen for always being there for me, and encouraging and supporting me. Last, but not least, the most important and beloved people in my life deserve the greatest thanks: my children Maria and Kristian and my husband Gert for endless love and support in everyday life.

Liv Marit Valen Schougaard, January 2020

Contents

LIST OF ORIGINAL PAPERS INCLUDING PUBLICATION STATUS 1

ABBREVIATIONS 2

TABLES 3

FIGURES 3

DEFINITIONS 4

INTRODUCTION 5

PATIENT-REPORTED OUTCOMES 6

Measurement properties of PRO measures 6
Administration of PRO measures 7

USE OF PATIENT-LEVEL PRO MEASURES IN CLINICAL PRACTICE 8

Applications of patient-level PRO measures in clinical practice 8
Evidence of the effect of using patient-level PRO measures in the clinical encounter 10
Evidence of the effect of using patient-level PRO measures in remote patient management 12

PATIENT-INITIATED FOLLOW-UP 15

PRO IN DENMARK 16

THE AMBUFLEX MODEL 16

The AmbuFlex history 16

AMBUFLEX/EPILEPSY 17

PRO-based follow-up in outpatients with epilepsy 18
Patient referral 19
The epilepsy questionnaire 19
The PRO-based algorithm 20
Review of responses 20

SYNTHESIS 21

AIMS 23

Hypotheses 24

METHODS 25

Literature search 27
Ethics 27
Statistical program 28
Study I: Determinants of referral to PRO-based follow-up 28
Study design and population 28
Questionnaire 28
Registers 28
Determinant variables 30
Outcome variable 31
Statistical analysis 32
Study II and Study III: Test-retest reliability of the PRO-based algorithm and WHO-5 33
Study design and population 33
Development of the epilepsy questionnaire and the PRO-based algorithm 33

The WHO-Five Well-Being Index 34
Statistical analysis 35
Study IV: The effectiveness of patient-initiated PRO-based follow-up 36
Study design and population 36
Recruitment and procedures 36
Randomization and blinding 37
Intervention 38
Outcome measures 39
Statistical analysis 40

RESULTS 43

Study I: Determinants of referral to PRO-based follow-up 45
Study population 45
Register-based determinants of referral to PRO-based follow-up 46
Questionnaire-based determinants of referral to PRO-based follow-up 46
Original data analyses and sensitivity analyses 46
Study II and Study III: Test-retest reliability of the PRO-based algorithm and WHO-5 47
Study population 47
Agreement and test-retest reliability of the PRO-based algorithm 47
Test-retest reliability of the items 49
Test-retest reliability and measurement error of the WHO-Five Well-Being Index 49
Study IV: The effectiveness of patient-initiated PRO-based follow-up 51
Study population 51
Use of healthcare resources (primary outcome) 52
Patient health outcomes and the patient perspective (secondary outcomes) 52
Attrition and sensitivity analyses 52
Process outcomes 53

DISCUSSION 55

Discussion of results 56
Study I: Determinants of referral to PRO-based follow-up 56
Study II: Test-retest reliability of the PRO-based algorithm 58
Study III: Test-retest reliability and measurement error of WHO-5 59
Study IV: The effectiveness of patient-initiated PRO-based follow-up 60
Critical evaluation of the methods 64
Study I: A prospective cohort study 64
Study II and Study III: Test-retest reliability studies 67
Study IV: A pragmatic randomized controlled study 68
Generalizability 72

CONCLUSION 73

PERSPECTIVE AND FUTURE RESEARCH 75

ENGLISH SUMMARY 79

DANSK RESUMÉ 81

REFERENCES 83

PAPERS

APPENDICES

LIST OF ORIGINAL PAPERS INCLUDING PUBLICATION STATUS

This dissertation is based on the following research papers:

I. Schougaard LMV, de Thurah A, Christensen J, Lomborg K, Maindal HT, Mejdahl CT, Vestergaard JM, Winding TN, Biering K, Hjollund NH. Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study. Qual Life Res 2020 Jan 3. (E-pub ahead of print).

II. Schougaard LMV, de Thurah A, Christiansen DH, Sidenius P, Hjollund NH. Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test-retest reliability study. BMJ Open 2018 Jul 25;8(7):e021337-2017-021337.

III. Schougaard LMV, de Thurah A, Bech P, Hjollund NH, Christiansen DH. Test-retest reliability and measurement error of the Danish WHO-5 Well-being Index in outpatients with epilepsy. Health Qual Life Outcomes 2018 Sep 6;16(1):175-018-1001-0.

IV. Schougaard LMV, Mejdahl CT, Petersen KH, Jessen A, de Thurah A, Sidenius P, Lomborg K, Hjollund NH. Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: study protocol for a pragmatic randomised controlled study. BMC Health Serv Res 2017 Jan 26;17(1):83-017-2015-8.

V. Schougaard LMV, Mejdahl CT, Christensen J, Lomborg K, Maindal HT, de Thurah A, Hjollund NH. Patient-initiated versus fixed-interval patient-reported outcome-based follow-up in outpatients with epilepsy: a pragmatic randomized controlled trial. J Patient Rep Outcomes 2019 Sep 13;3(1):61-019-0151-0.

ABBREVIATIONS

BI	Hospital Business Intelligence Register
CI	Confidence Interval
CONSORT	Consolidated Standard of Reporting Trial
COSMIN	Consensus-based standards for the selection of health measurement instruments
CPR	Personal identification number
DNPR	Danish National Patient Register
DREAM	Danish Register for Evaluation and Marginalization
ePRO	Electronic patient-reported outcome
GSES	General Self-Efficacy Scale
HLQ	Health Literacy Questionnaire
HRQOL	Health-related quality of life
ICC	Intraclass correlation coefficient
ICD-10	International Classification of Diseases, version 10
IQR	Interquartile range
ISOQOL	International Society for Quality of Life Research
ISPOR	International Society for Pharmacoeconomics and Outcome Research
ITT	Intention-to-treat
MAR	Missing at random
MDC	Minimal detectable change
MI	Multiple imputation
MIC	Minimal important change
MNAR	Missing not at random
NASSS	Nonadoption, abandonment, scale-up, spread, and sustainability
PAM	Patient Activation Measure
PRO	Patient-reported outcome
QUIPS	Quality In Prognosis Studies
RCT	Randomized controlled trial
RR	Risk ratio
SCL-92	Symptom Checklist 92
SD	Standard deviation
SF-36	Short Form 36 Health Survey
WHO-5	World Health Organization-Five Well-Being Index

TABLES

Table 1 Summary of potential applications of individual use of patient-level PRO data
Table 2 Overview of randomized controlled trials regarding use of PRO measures in remote patient management
Table 3 Number of outpatients with epilepsy in the Central Denmark Region
Table 4 Distribution of the PRO-based algorithm in outpatients with epilepsy in the Central Denmark Region from 2012 to 2020, *n*(%)
Table 5 Overview of study design methodology in the four studies
Table 6 Overview of determinant variables and register and questionnaire data sources (120)
Table 7 Secondary self-reported outcomes, data measures, and scoring
Table 8 Overview of study population characteristics in the four studies
Table 9 Number of patients who had been referred to PRO-based follow-up, terminated outpatient care, or died at the three time points used in the study, *N*=802 *
Table 10 Agreement between the automated PRO-based algorithms from test 1 to test 2 (142)
Table 11 Healthcare utilization during an 18-month follow-up period (144)
Table 12 Patient-reported outcomes measured 18 months after randomization (144)

FIGURES

Figure 1 Overview of measurement properties according to the COSMIN taxonomy (23,25)
Figure 2 Distribution of questionnaires received by the WestChronic/AmbuFlex system according to use at either group or patient level; updated from (75).
Figure 3 Overview of the pathway of PRO-based follow-up among outpatients with epilepsy
Figure 4 The PRO overview shown to the clinicians in the Electronic Health Record (4)
Figure 5 Aims of the studies according to the PRO-based follow-up model
Figure 6 Overview of study populations in the four studies
Figure 7 The unidimensional structure of the WHO-Five Well-Being Index (WHO-5)
Figure 8 Overview of study design and procedures in Study IV
Figure 9 The patient-initiated/open access website 'My Epilepsy' (138)
Figure 10 Flowchart of patients included in the study (120)
Figure 11 Flowchart of response method in test 1, randomization of response method in test 2, non-responders in test 2, and number of patients included in the analyses (142)
Figure 12 Test-retest reliability from test 1 to test 2 of the pooled PRO-based algorithms (*n* = 554), web-web (*n* = 166), paper-paper (*n* = 112), and the mixed groups (web-paper or paper-web, *n* = 276) (142)
Figure 13a Bland-Altman plot of differences in WHO-5 score between test 1 and test 2 plotted against the mean, *N* = 540 (143)
Figure 13b Differences in the WHO-5 score between test 1 and test 2 plotted against the mean in the four methods of administration groups: web-web (*n* = 164), paper-paper (*n* = 107), web-paper (*n* = 233), and paper-web (*n* = 36) (143)
Figure 14 The study CONSORT flow diagram (144)
Figure 15 The non-adoption, abandonment, scale-up, spread, and sustainability framework from Greenhalgh et al. (179)

DEFINITIONS

Patient-reported outcome (PRO)	A measurement of any aspect of a patient's health that comes directly from the patient without interpretation of the patient's response by a clinician or anyone else (1,2).
Patient-reported outcome measure	A questionnaire that measures patients' perceptions of the impact of a condition and its treatment on their health (3).
Questionnaire	A term often used to describe a patient-reported outcome or other collection of self-reported items (2).
Electronic patient-reported outcome measure	A PRO measure that is completed electronically (3).
Remote PRO-based follow-up or telePRO	PRO measures are used as the basis for follow-up in which scheduled questionnaires replace in-clinic visits. A red flag approach/PRO-based algorithm is used to identify patients who need or wish clinical attention (4).
PRO-based algorithm	PRO measures are divided into three colors: green, yellow, or red based on what the clinicians consider clinically important to react on to identify patients who need attention (4).
PRO overview	A graphical overview displaying PRO measures. Clinicians use the interface to guide clinical decision-making (4).
Remote PRO-based monitoring *	PRO measures are used to monitor symptoms distress between scheduled in-clinic visits.
Patient-initiated follow up	Follow-up activities that allow patients to initiate hospital outpatient follow-up appointments on an "as required" basis compared with the traditional "regular in-clinic appointment" model. The main principle is to reduce inappropriate regular follow-up appointments at times when patients are feeling well and symptom-free (5,6).
Remote patient-initiated PRO-based follow-up or open access telePRO *	PRO measures are used as the basis for follow-up by which non-scheduled questionnaires initiated by the patients replace scheduled in-clinic visits.

* Ad hoc definition used in this PhD dissertation

INTRODUCTION

The number of people with long-term chronic conditions is increasing worldwide (7). This is also the case in Denmark, as chronic conditions involve two-thirds of the adult population (8,9). This increase puts pressure on the healthcare system to manage the balance between rapid acute care management and the needs of patients with chronic conditions (10). Moreover, the healthcare system focuses on the goal of involving patients' perspectives on their own health and healthcare delivery to improve quality of care (11). The patient perspective can only be evaluated by the patients themselves; hence, the use of patient-reported outcome (PRO) measures has gained attention during the last decade at all levels of the healthcare system (3,12). These outcomes are of major importance as the ultimate goal in healthcare is to reduce symptoms and improve patients' health status as well as how they function in daily life (13). PRO measures have the potential to facilitate patient-centered care through different types of use (14-16) and could contribute to the reorganization of healthcare delivery to patients with chronic conditions (17).

Group-level PRO data could have a role in organizational use to evaluate quality of care, the performance of healthcare providers or organizations, population monitoring, and in value-based healthcare delivery (3,18). Group-level PRO data could also be used at the individual patient level in clinical practice in which PRO data may have a role in shared decision-making and individual prognosis on functioning and symptoms (14,16). Individual patient-level PRO data could be used in the clinical encounter between patients and clinicians to guide the conversation and to focus on issues reported as important by the patients themselves (3,16,19). In addition, individual patient-level PRO data can be used to flag the need for further clinical attention in remote patient management (14,17). This could support cost-saving activities and reduce unnecessary outpatient visits for stable patients with long-term conditions and perhaps also enhance patient-centered care (17). This dissertation focuses on the individual use of patient-level PRO data in remote outpatient follow-up.

PATIENT-REPORTED OUTCOMES

Patient-reported outcome (PRO) is defined as: *"any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else*" by the US Food and Drug Administration (1). Thus, PRO is an umbrella term that covers a broad range of different constructs, such as, symptoms; physical, social, and mental function; health perceptions; and health-related quality of life (HRQOL) (20,21). These aspects encompass different levels of patient health status or outcomes and are often non-observable subjective aspects (e.g. pain, fatigue, and depression) that only can be assessed by the patients themselves. Since these aspects cannot be observed directly, they need to be operationalized to be measured. Non-observable constructs can be measured in a standardized questionnaire, also called a PRO measure (19). A PRO measure includes specific domains relevant to the patient's condition and health status (22). Each domain includes one or more items that reflect the construct to be measured. The items are scored and often based on a procedure to calculate a summary score. A PRO measure can be generic, domain, or disease specific (22). PRO measures can play a central role in clinical practice and health research as they can be used in diagnosis, prognosis, and evaluation of the effects of a medical treatment or intervention (22).

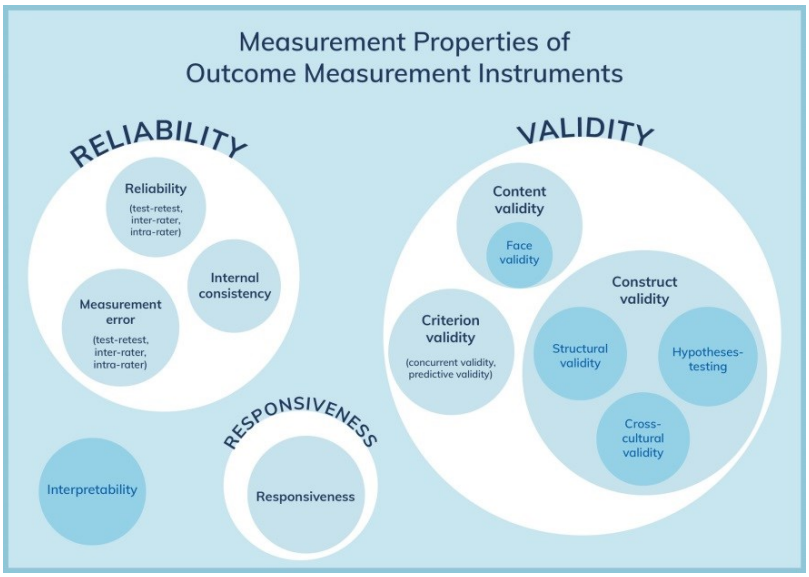
Measurement properties of PRO measures

A PRO measure must be used adequately according to its purpose, contain valid and reliable information, and users must know how to interpret the results a PRO measure produces (22). In both the development and evaluation of a PRO measure, measurement properties such as validity and reliability must be considered. The variation in terminology and definitions can be confusing in the literature; therefore, an international Delphi study was accomplished in 2010 to achieve consensus based standards for the selection of health measurement instruments (COSMIN) (22-24). COSMINs terminology is used throughout this dissertation.

Figure 1 shows the COSMIN taxonomy, which includes three main measurement properties: reliability, validity, and responsiveness (23,25). Reliability relates to how consistently patients respond to a PRO measure over time, for example, if the PRO measure was measured again one week later, would it show the same results if the patient's health status did not change? Further, to what degree are the repeated PRO measures free from measurement error? (23). This can be tested in a test-retest reliability design. The validity of PRO measures refers to whether they measure what they are intended to measure. Validity can be divided into several measurement properties; two that are highly relevant to PRO measures are content and construct validity. Content validity relates to whether all aspects of the construct of interest are covered in the PRO measure and whether the target population finds the PRO measure relevant and comprehensible (face validity) (23). Construct validity relates to how well the PRO measure performs compared to other PRO measures that aim to measure the same construct (23). Criterion validity could also be relevant and

relates to the degree to which a PRO measure is an adequate reflection of a measurement gold standard (23); however, often a gold standard is not available. The responsiveness of a PRO measure relates to whether it can detect a change in the measured construct over time (23). This is a relevant measurement property if the PRO measure is intended for evaluation of an effect of an intervention.

Figure 1 Overview of measurement properties according to the COSMIN taxonomy (23,25)



Administration of PRO measures

Traditionally, PRO measures have been developed for paper-and-pencil administration. However, electronic data collection of PRO measures, termed ePRO, are more commonly used since the use of electronic devices, such as computers, tablets, and smart phones has become more widespread, not only in everyday life, but also to capture data in the healthcare system and health research (26). Several advantages of using ePRO data collection have been pointed out, such as less administrative burden, automatic scoring of data, immediately accessible data, and more accurate and complete data (27,28). In the PRO Guidance document, the US Food and Drug Administration states that the migration of paper-based PRO measures to a web-based PRO version should be supported by evidence (1). On the basis of this statement, the International Society for Pharmacoeconomics and Outcome Research (ISPOR) has developed recommendations to ensure comparability of electronic and paper-based PRO measures (29). The recommendations are divided into three levels based on the magnitude of modifications of the paper-based version needed to convert it to a web-based version for use in an integrated electronic platform. In the literature, four comprehensive systematic reviews and meta-analyses including studies from 1991 to 2015 have examined the relationship between web- and paper-based methods of PRO administration. The results indicate that web- and paper-administered PRO measures are equivalent (28,30-32).

USE OF PATIENT-LEVEL PRO MEASURES IN CLINICAL PRACTICE

PRO measures are increasingly being used in clinical practice and have attracted more attention from health authorities and decision-makers in the healthcare system during the last decade (14). The value of integrating the patient perspective into healthcare is considered to be of major importance (11). PRO measures could be considered as a means of achieving the goal of involving patients in the healthcare process (3,12,14). Individual use of group-level PRO data involves shared decision-making or individual prognosis on functioning or symptoms (14). As the focus of this dissertation is the individual use of patient-level PRO data, the use of group-level PRO data at the individual patient-level will not be further elaborated on.

Applications of patient-level PRO measures in clinical practice

In 1992, the three significant applications of health status information in clinical practice were screening, decision-making regarding care, and monitoring the effect of care (33). Today, PRO measures are applied for a broad range of purposes (14,19,34). However, defining different applications of individual patient-level PRO measures in clinical practice is not straightforward. This is because the intervention and implementation of PRO into clinical practice is complex and far from a "one size fits all" solution (34). Moreover, often multiple applications of PRO measures occur in clinical practice.

The majority of theoretical and empirical explorations on individual patient-level PRO data focus on the clinical encounter between a patient and a clinician. In this case, the patient responds to PRO measures at home or in the clinic before an in-clinic visit. In 2008, Greenhalgh et al. outlined a taxonomy of different applications for using PRO measures in clinical practice (16). Others have described similar applications (3,19,34,35). They all point out three main applications if the aim of individual patient-level PRO data is to improve care at the patient-clinician interface. Firstly, PRO data can be used as a screening tool to detect physical, social, or mental problems. Secondly, PRO data can be used as a monitoring tool to identify treatment effects. In both cases, the PRO data can be used to support clinical decision-making, whether to stop, modify, or add plans for treatment and supportive care, e.g. diagnosis and referral. Thirdly, PRO data can be used as a means to establish patient-centered care by fostering a prompt discussion between patients and clinicians regarding issues reported as important by the patient. Hence, patients may become more involved in decisions about their care, which may increase their ability to manage their own health. This could contribute to increased patient satisfaction, self-management, and improved health outcomes (16). This mechanism is supported by a theoretical framework reported by Santana et al. (36).

Use of PRO measures in the clinical encounter is added to the current work flow in which patient care is organized around visits. In 2008, Donaldson outlined that other means of contact can be valuable and that PRO measures in the future could be used as a way to organize healthcare

delivery (37). If PRO measures are used to help organize the care process, the PRO measures should be used at the center of healthcare, and not an added task to usual care (37).

In 2014, Hjollund et al. described the use of PRO measures as the basis for remote outpatient follow-up in Denmark (17). And in 2017, Basch described the use of PRO measures in remote patient monitoring in which PRO measures are captured electronically between in-clinic appointments (38). In 2019, both applications were included in a paper by Calvert et al. (14). Although remote management has been used for decades in telemedicine solutions in the healthcare system, the use of PRO measures in remote patient management is a relatively new initiative in clinical practice (17). This allows new opportunities such as real-time monitoring of symptoms, flexible scheduling of hospital appointments, early detection of problems, and a prompt clinical intervention if needed (14). In such cases, PRO measures could contribute to the reorganization of the healthcare system for follow-up activities in patients with long-term conditions by prioritizing or optimizing the use of healthcare resources and promoting patient-centered care. However, the applications of using PRO measures in remote patient management depend on the purpose of the PRO intervention and perhaps also the disease complexity.

In this dissertation, the following terms are used: *remote PRO-based monitoring*, *remote PRO-based follow-up*, and *remote patient-initiated PRO-based follow-up*.

In *remote PRO-based monitoring*, PRO measures can facilitate closer monitoring of symptom distress between scheduled in-clinic visits. If real time monitoring of symptoms is fed back to clinicians using alerts, a prompt clinical intervention could be performed and progression of disease activity may be avoided (38). Remote PRO-based monitoring could be used in the management of complex conditions with fluctuating disease activity, for example, cancer, heart failure, renal failure, and chronic obstructive pulmonary disease.

In *remote PRO-based follow-up*, PRO measures are used as the basis for follow-up, scheduled questionnaires replacing in-clinic visits. In this case, PRO measures could be used to identify patients who need clinical attention by using red flag alerts to clinicians. This enables a more tailored and flexible individual follow-up activity, and unnecessary visits can be avoided (17). Remote PRO-based follow-up could be used in the management of stable chronic conditions, which is often the case for patients with, for example, asthma, diabetes, rheumatoid arthritis, and epilepsy.

In *remote patient-initiated PRO-based follow-up*, PRO measures can be used as the basis for follow-up, non-scheduled questionnaires initiated by the patients replacing in-clinic visits. In patient-initiated follow-up, patients are not routinely seen in hospital, but are instructed to contact the hospital if a real or perceived clinical event arises (5). This model could potentially provide patients with more control over their own health and may increase patients' self-management concurrent with better utilization of healthcare resources. Using PRO measures in remote patient-initiated follow-up must be considered as a theoretical application, as I have not been able to identify studies that describe

usage of PRO-based patient-initiated follow-up, but only patient-initiated follow-up using telephone contact as the main approach for patients to initiate contact to the healthcare system (5). The evidence related to patient-initiated follow-up is described in a separate section.

Based on different applications described in the literature (3,14,16,17,19,34,35,38) and patient-initiated follow-up, a summary of potential applications of individual use of patient-level PRO data in clinical practice is presented in Table 1.

Table 1 Summary of potential applications of individual use of patient-level PRO data

Clinical setting	Purpose	Description
Patient-clinician encounter	Screening	PRO data help identify undetected problems.
	Monitoring	Repeated PRO data help track progress over time and/or evaluate treatment effect over time.
	Patient-centered care	Review of PRO data helps improve patient-clinician communication by prioritizing and addressing problems or concerns important to the patients. PRO data increase patients' understanding of symptoms and disease. PRO data increase patients' ability to manage their own health.
Remote patient management	Prioritize and/or optimize use of healthcare resources and promote patient-centered care	**PRO-based monitoring:** PRO data facilitate closer monitoring of disease activity and prompt clinical interventions to prevent disease progression between scheduled in-clinic visits.
		PRO-based follow-up: PRO data help promote individual flexible scheduling of follow-up appointments by which scheduled questionnaires replace in-clinic visits.
		Patient-initiated PRO-based follow-up: PRO data help promote individual flexible scheduling of follow-up appointments by which non-scheduled questionnaires initiated by the patients replace in-clinic visits.

References (3,14,16,17,19,34,35,38)

Evidence of the effect of using patient-level PRO measures in the clinical encounter

The evidence on this topic is comprehensive; hence, the summary of the evidence is based on secondary literature including systematic or literature reviews identified after a systematic literature search (Appendix 1) and by reviewing relevant references.

Several reviews have investigated the effect of using patient-level PRO measures in clinical practice in which PRO measures are collected from patients before a consultation, with feedback to the healthcare providers who aim to use the PRO measures in the clinical encounter (39-50). In addition, two upcoming reviews have recently been presented at the 26th annual conference of the International Society for Quality of Life Research (ISOQOL) in October 2019 (51). The included reviews are primarily based on randomized controlled trials (RCT) in various patient populations and clinical settings. The measured effects of using PRO measures in the clinical encounter are traditionally divided into three outcome categories: process of care, patient health outcomes, and

patient satisfaction. In the present summary, a fourth outcome category regarding use of healthcare services has been included. The process of care was related to how care was delivered, for example, the quantity and quality of patient-clinician communication, problem-detection, referral and treatment rates, and patients' adherence to treatment. Patient health outcomes involved, for example, changes in patients' HRQOL, symptoms, well-being, functioning, or psychological distress.

The effect of using PRO measures in the clinical encounter was often related to the process of care. Improved patient-clinician communication has been reported in nine reviews (39,41,43,44,46-50). PRO measures were found effective in prompting discussion of troublesome symptoms, which made it possible to focus the conversation on issues relevant to the patient. Five reviews found that use of PRO measures increased the detection of problems; the detection of mental issues was particularly improved (43,44,46,47,49). In addition, four reviews reported improved changes in patient management; for example, increase in the rates of diagnosis and higher referral rates to other professionals have been reported in four reviews (43,47-49). In an unpublished review presented at the annual ISOQOL conference in October 2019, Sidey-Gibbons et al. stated that the most profound finding was related to identifying health issues leading to appropriate diagnosis and referral (51). Increased patient satisfaction in the PRO intervention group has been reported in four reviews (41,43,46,48). Clinician acceptability has not been the focus of many of the studies included in the reviews. Only one review reported that clinician acceptability regarding managing and enhancing care was moderate to high, and that nurses were more positive (41). Similarly, in another review, most of the clinicians considered the feedback to be useful (48). Furthermore, one review reported that the PRO interventions improved clinicians' adherence to recommended clinical practice (40). The impact of using PRO measures in the clinical encounter on patient health outcomes and healthcare service outcomes is more uncertain. A review from 2018, including 18 RCTs with health outcomes as primary outcome, found only three studies with robust positive effects, such as improvement in HRQOL and psychosocial health (39). The other studies reported either non-robust effects or no significant differences (39). Similarly, in other reviews only few positive effects on patient health outcomes was found (41-47,49). Few studies have evaluated the use of healthcare services when using PRO measures in the clinical encounter. One review included five studies that assessed the number of patients making use of healthcare services and the frequency of contact with clinicians; however, the results are conflicting (41). Cost-effectiveness evaluations related to this topic were not identified.

Heterogeneity in terms of PRO interventions, outcome measures, and the methodological quality of the included studies is described as a general weakness in all reviews. The inconclusive impact of PRO interventions on patient health outcomes could be related to the focus on distal outcomes, without understanding the effect upon proximal outcomes, such as how the PRO intervention operates between patients and clinicians in clinical practice (19).

Evidence of the effect of using patient-level PRO measures in remote patient management

The evidence on this topic is scant; hence, the summary of the evidence is based on primary literature including identified RCT studies after a systematic literature search (Appendix 1) and by reviewing relevant references.

A study was included if it evaluated a remote PRO-based intervention that either replaced in-clinic visits or was used to monitor symptoms between in-clinic visits. It was also a criterion that the PRO data were fed back to clinicians, but there were no restrictions on types of PRO measures used, health condition, or the setting or country in which the study was conducted. Feasibility or pilot RCTs were not included. The measured effect of using PRO measures in remote follow-up or monitoring was also divided into four outcome categories: process of care, patient health outcomes, patient satisfaction, and use of healthcare services.

Four studies were identified that investigated the effect of a PRO-based intervention in remote follow-up in which PRO data replaced in-clinic visits (52-55). However, in three of the studies, the PRO-based interventions were supplemented with in-clinic visits; thus, only some of the in-clinic visits were replaced by PRO measures during follow-up (52,53,55). A Danish non-inferiority RCT study examined the effect of a PRO-based intervention in patients with stable rheumatoid arthritis (52). As hypothesized, the study found no differences in disease activity between the study groups. The study showed no differences in HRQOL and self-efficacy in the PRO-based interventions compared to traditional follow-up with scheduled in-clinic visits. However, the study showed that patients in the PRO-based intervention groups had fewer outpatient visits (52). A RCT study from the Netherlands examined the effect of a PRO-based telemedicine intervention in patients with inflammatory bowel disease (53). The study showed that PRO-based follow-up resulted in fewer outpatient contacts and hospital admissions, and improved medicine adherence compared to traditional in-clinic follow-up. The study found no differences in patient health outcomes such as disease progression, HRQOL, and self-efficacy. Furthermore, no difference was found in patient satisfaction (53). A Canadian RCT study in patients undergoing ambulatory breast reconstruction showed that patients that used a remote PRO-based mobile application had fewer in-clinic visits than patients with scheduled in-clinic follow-up during the first 30 days after the operation (54). The study found no differences in complication rates or satisfaction between the groups, but the PRO intervention group reported a higher convenience score than the control group (54). A RCT study from the UK investigated the use of remote ePRO for early rheumatoid arthritis management (55). The study found no differences in disease activity after 3, 6, and 12 months of management, and the patients' adherence to treatment was statistically significantly higher in the PRO intervention group than the control group after 12 months of treatment (55).

Four studies which investigated the effect of a PRO-based intervention in remote monitoring in which PRO data were used to monitor symptoms between in-clinic visits were identified. A French RCT study investigated the effect of weekly PRO monitoring in patients with lung cancer in order

to prevent disease progression. The study showed an increase in survival of patients who were PRO monitored compared to patients in traditional follow-up (56,57). In addition, the PRO intervention was considered cost-effective (58). Furthermore, a RCT study from the US has shown positive effects of monitoring symptoms during chemotherapy in patients with metastatic cancer (59). The patients in the PRO-based monitoring group had an improvement in HRQOL after 6 months and increased survival compared to patients who received usual care (59,60). A RCT study from the US examined web-based symptom management after treatment for breast cancer (61). The study evaluated its impact on use of healthcare resources; however, no differences compared to standard care were reported (61). Furthermore, another RCT study from the US showed reductions in symptom burden when symptoms were monitored daily using an interactive voice system in patients undergoing chemotherapy treatment (62). Following this study, a comprehensive PRO system has been developed to track and respond to PRO data between clinic visits (63).

The included RCT studies in this section are presented in Table 2. In summary, the research indicates that remote follow-up based on PRO measures maintain quality of care and contribute to lower use of healthcare services in patients with long-term conditions. For patients with complex health conditions who may have a risk of a high degree of symptom burden, frequent PRO monitoring between in-clinic visits indicates positive effects on patients' health outcomes such as HRQOL, symptom burden, and survival. Limitations of the studies are mainly related to lack of blinding, missing outcome data, measurement of self-reported outcome data, or lack of generalizability due to a selected patient group or single-center study. Further research is important to gain deeper insight into the impact of using PRO measures in remote patient management. Several RCT protocols related to this topic have been identified in both patients with kidney disease and in oncology patient populations (64-68).

Table 2 Overview of randomized controlled trials regarding use of PRO measures in remote patient management

Publication ID Country	**Setting (e.g. primary care, outpatient care)**	**Healthcare providers**	**Patient population *n* (Intervention versus control)**	**Intervention**	**Type of PRO measure**	**Method of administration**	***Primary* and selected secondary outcomes**
de Thurah (52), 2018 Denmark	Outpatient, hospital-based	Rheumatologists Rheumatology nurses	Rheumatoid arthritis (RA), stable disease activity Intervention (nurse) = 97 Intervention (doctor) = 99 Control = 98	Online or paper PRO response prior to scheduled telephone consultation every 3–4 months combined with one yearly in-clinic visit	The Flare-RA instrument	Self-administered Web or paper	*Disease activity* Level of function Quality of Life Self-efficacy
de Jong (53), 2017 Netherland	Outpatient, hospital-based	Gastroenterologist Nurses	Inflammatory bowel disease (IBD) Intervention = 465 Control = 444	Online PRO response at least every 3rd month combined with one yearly in-clinic visit	The Monitor IBD At Home (MIAH) questionnaire	Self-administered Web	*Use of healthcare resources* *Quality of care* Self-efficacy Adherence to treatment
Armstrong (54), 2017 Canada	Outpatient, hospital-based	Physicians	Patients undergoing elective breast reconstruction Intervention = 34 Control = 36	Daily online PRO responses for 2 weeks and then weekly for the remaining 2 weeks	Pain visual analog scale and 9-item quality of recovery questionnaire	Self-administered Web	*Use of healthcare resources* Satisfaction Convenience Adverse events
Denis (56), 2017 France	Outpatient, cancer clinic	Oncologists Nurses	Lung cancer Intervention = 67 Control = 66	Weekly online PRO monitoring combined with in-clinic visits every 3rd month	12 symptom items	Self-administered Web	*Overall survival*
Mooney (62), 2017 USA	Outpatient, cancer clinic	Oncologists Nurses	Patients receiving cancer chemotherapy Intervention = 180 Control = 178	Patients called the IVR-system daily reporting severity of symptoms	11 symptom items	Self-administered Phone	*Overall symptom severity* Number of symptom days
Basch (59), 2016 USA	Outpatient, cancer clinic	Oncologists Nurses	Metastatic cancer (computer-experienced) Intervention = 286 Control = 253	Weekly online PRO monitoring between scheduled in-clinic visits during chemotherapy	12 common symptom items	Self-administered Web	*Change in quality of life* Emergency room visits and hospitalization Survival
El Miedany (55), 2016 UK	Rheumatology centers	Rheumatologists Rheumatology nurses	Rheumatoid arthritis, early stage Intervention = 112 Control = 112	Monthly ePRO questionnaires between scheduled in-clinic visits every 3rd month	A PRO instrument including 11 domains	Self-administered Web	*Disease activity* Adherence to treatment
Wheelock (61), 2015 USA	Outpatient, cancer clinic	Oncologists Nurses	Breast cancer Intervention = 59 Control = 41	Online PRO response every 3rd month combined with 3 in-clinic visits during 18 mo.	SF-36 PHQ-8 Symptom items	Self-administered Web	*Time between symptom reporting and remote evaluation of symptoms* Use of healthcare resources

Abbreviations PRO: patient-reported outcome; IVR: interactive voice response; SF-36: Short Form 36 Health Survey; PHQ-8: 8 item Personal Health Questionnaire Depression Scale

PATIENT-INITIATED FOLLOW-UP

I did not identify any studies that investigated the effect of a patient-initiated intervention in which PRO measures were used to initiate contact with the healthcare system. Hence, the literature search was expanded to include studies investigating the effect of a patient-initiated intervention in which patients initiated contact with the healthcare system by other methods, such as by telephone. The summary of the evidence is based on both the primary and secondary literature after a systematic literature search (Appendix 1) and by reviewing relevant references.

A mixed methods study from the UK published in 2019 examined experiences with patient-initiated follow-up in patients with endometrial cancer. The study found that the women who were offered patient-initiated follow-up contacted the clinic more often during the first 6 months compared to the second 6 months. Qualitative data indicated that some of the women perceived that they had more control over their own health, but others reported that they missed a more personal face-to-face contact (69). A Danish RCT study from 2018 also concerning patients with endometrial cancer (early stage) found that patient-initiated follow-up increased "Fear of Cancer Recurrence Inventory", and that patients had fewer examinations at the department compared to traditional, scheduled in-clinic follow-up. The study concluded that the use of patient-initiated follow-up should balance benefits and harms (70). In a prospective cohort study from the UK published in 2017, in patients with curatively treated colorectal cancer no differences were seen in satisfaction of patients with patient-initiated follow-up compared to patients with traditional outpatient follow-up. Furthermore, healthcare service costs were found to be higher in patient-initiated follow-up due to a self-management education program (71). Additionally, a British RCT study from 2016 in patients with breast cancer found no differences in HRQOL, depression, or anxiety between patient-initiated follow-up and traditional follow-up (72).

We identified three systematic reviews, including RCT studies from 1980 to 2013, that investigated the effect of patient-initiated follow-up (5,6,73). The studies involved patients with breast cancer, rheumatoid arthritis, and inflammatory bowel disease. Overall, the included studies did not find any differences in patient health outcomes, such as disease activity, HRQOL, and self-efficacy, between patient-initiated follow-up and traditional follow-up. However, better patient satisfaction and less use of healthcare services in patient-initiated follow-up have been reported by some studies (5,6,73).

In summary, the evidence is inconclusive and based on studies with methodological limitations. Despite effects indicating less use of healthcare resources and increased patient satisfaction in some studies, no other effects have been documented. Furthermore,

we lack evidence on the effect of using a patient-initiated intervention combined with PRO measures as the basis for follow-up in outpatient care.

PRO IN DENMARK

The Danish Health Authority has established several ambitious initiatives to expand the clinical use of PRO measures in the healthcare system nationally (8,74). For example, in 2016, in the national budget negotiations between the government, the regions, and the municipalities, agreement was reached that implementation of PRO measures should be initiated in hospital departments before 2020 in three disease areas: epilepsy, prostate cancer, and breast cancer (8). This agreement was partly based on experiences with use of PRO measures in the Danish generic PRO system, AmbuFlex (4,17).

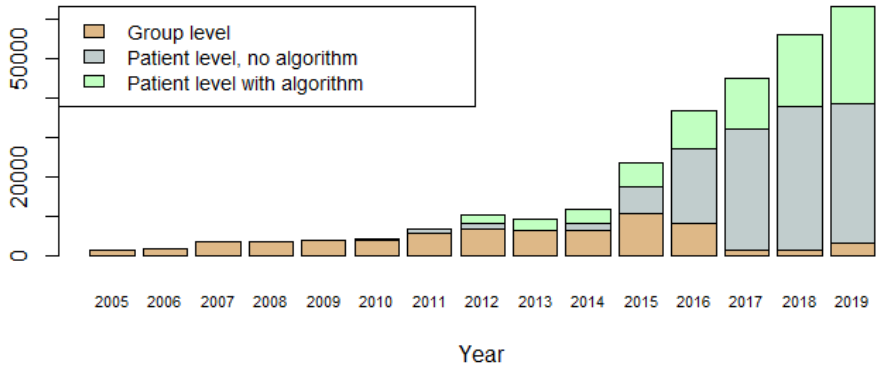
THE AMBUFLEX MODEL

In AmbuFlex, PRO measures are used as the basis for remote follow-up instead of fixed appointments in many chronic and malignant diseases (4). This is termed *PRO-based follow-up* or *telePRO*. Typically, PRO data are captured at patients' homes and used to flag whether a patient wants attention and used by clinicians to decide whether patients need clinical attention. The overall aim is to achieve patient-centered care, improve quality of care, and optimize the use of resources in the healthcare system (4). The AmbuFlex model consist of three generic elements including PRO data collection, a PRO-based algorithm, and a PRO-based graphical overview (4,17). AmbuFlex has been used in clinical practice since 2012, and as of January 2020, AmbuFlex has been implemented in 35 different patient groups in hospitals in three of the five Danish Regions.

The AmbuFlex history

In 2004, the generic PRO system WestChronic was developed for research purposes of group-level PRO data in longitudinal epidemiological studies (17,75). In 2007, it was decided to build on the positive experiences with the system in relation to feasibility and high response rates and to further develop the system to use PRO data at the individual patient level in clinical practice (75). This system was named AmbuFlex. Figure 2 illustrates the development from group-level projects to individual-level use of PRO data from 2005 to 2020 (75). Patient-level use of PRO data has nearly completely taken over since 2017.

Figure 2 Distribution of questionnaires received by the WestChronic/AmbuFlex system according to use at either group or patient level; updated from (75)



The first version of the AmbuFlex system was tested among patients with heart failure in 2009 (17). In 2011, it was decided to develop and implement a PRO solution for outpatients with epilepsy using the AmbuFlex system (4,17). During the first year, nearly 2000 patients were referred and the use of PRO data combined with a PRO-based algorithm indicated that 48% of the patients did not need clinical attention (4). In 2013, it was decided to implement three new AmbuFlex PRO solutions yearly in the Central Denmark Region. In 2017, a joint steering group across all five Danish Regions was established, and since then AmbuFlex has taken part in the national work and strategy regarding use of PRO data at the individual patient level in Denmark.

AMBUFLEX/EPILEPSY

Epilepsy is a long-term condition with a prevalence of approximately 0.5–1% in the general population (76,77). The incidence rate varies with age, ranging from a high level in children to a low level between 20 and 40 years and thereafter a gradual increase is seen (76). Epilepsy is characterized by recurrent seizures affecting both physiological and socio-psychological aspects of life (78,79). These aspects can only be assessed by the patients themselves and could be measured using PRO measures in the care of people with epilepsy.

Nixon et al. identified 26 epilepsy-specific PRO measures that have primarily been used at the group level in epilepsy research (80), and other sources have also been identified (81-83). The use of PRO measures at the patient level in patients with epilepsy was first described in 1995 by Wagner et al. (84). They pointed out that health status information could reveal decline or improvement in patients' physical and mental functioning and provide relevant information to the clinicians (84). However, in the following years, to my

knowledge, only two minor studies regarding use of PRO measures in a clinical setting involving patients with epilepsy have been published (85,86), and these studies do not describe use of PRO measures in remote outpatient follow-up.

PRO-based follow-up in outpatients with epilepsy

A PRO solution was developed and implemented for outpatients with epilepsy in three neurological departments in the Central Denmark Region in 2012. Before implementation, follow-up was based on in-clinic visits or telephone consultations. The rationale for developing the solution was based on an increased number of outpatients with stable disease activity. Nonetheless, the need for monitoring patients and identifying potential issues remained important. The use of PRO measures in remote follow-up was driven by clinicians who worked with this patient group. The overall aim was to improve quality of care by facilitating greater flexibility in individual care, promoting patient-centered care, and achieving better utilization of healthcare resources (4).

In 2016, based on the national agreement by the health authorities, the process of implementing the solution in the other Regions started. A standardized epilepsy questionnaire and PRO-based algorithm is used in all five Danish Regions, three of which use the AmbuFlex system.

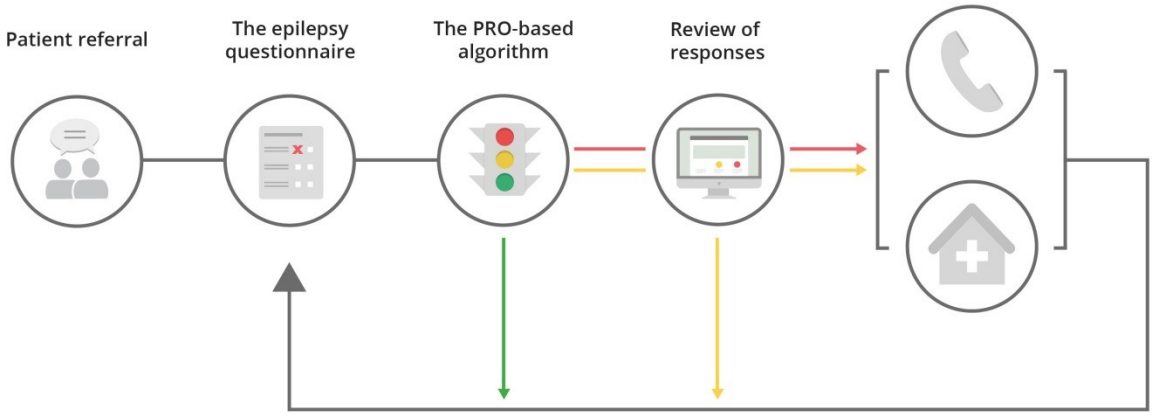
In the Central Denmark Region, as of January 2020, a total of 3062 epilepsy outpatients were attending PRO-based follow-up (Table 3), which is approximately 50% of the entire population of patients with epilepsy. Patients who do not attend PRO-based follow-up receive conventional follow-up including in-clinic visits or telephone follow-up. Figure 3 illustrates the pathway of using PRO-based follow-up in outpatients with epilepsy. The different elements are described in the following sections.

Table 3 Number of outpatients with epilepsy in the Central Denmark Region

	Aarhus University Hospital	**Regional Hospital Holstebro**	**Regional Hospital Viborg**	**Total**
Outpatients with epilepsy as of Nov. 2019 *	4000	1300	600	5900
Patients referred to PRO-based follow-up from 2011 to 2020	4251	1026	462	5739
Patients attending PRO-based follow-up as of January 2020	2100	611	351	3062

* Approximately number of patients based on information provided by the departments

Figure 3 Overview of the pathway of PRO-based follow-up among outpatients with epilepsy



Patient referral

Patients attending PRO-based follow-up are individually referred. Referral is managed in daily clinical practice, and participation is determined by a clinician together with the patient's own preferences and willingness to participate (4). Patients are referred to an individual fixed questionnaire interval, which is either 3, 6, 12, or 24 months. Thus, we also use the term *fixed-interval PRO-based follow-up*. No public standardized guidelines related to referral to PRO-based follow-up have been developed for outpatients with epilepsy. However, the solution is targeted to patients ≥ 15 years old with no cognitive disabilities. In addition, the patients must be able to read and understand Danish, as the questionnaire and patient information are only available in Danish. Other aspects such as co-morbidity and health literacy might also be relevant for clinicians to consider before deciding whether to refer a patient to PRO-based follow-up.

The epilepsy questionnaire

Patients complete a disease-specific questionnaire at the pre-defined individual intervals. If possible, the AmbuFlex system prompts patients through "e-Boks" (secure Danish personal e-mail platform) to fill in the questionnaire, otherwise, it is sent via surface mail. Thus, patients can complete either a web or paper version of the questionnaire. Up to three reminders are sent if the patients do not respond. The patients have access to their own questionnaire responses via secure login to the Danish eHealth Portal (https:\\Sundhed.dk).

The questionnaire includes 47 items covering several topics related to epilepsy, e.g. number of seizures, medicine adherence, symptoms, general health, and psychosocial function (Appendix 2). These aspects are measured using established PRO measures such as the WHO-Five Well-Being Index (WHO-5) (87,88), items from the Short Form 36 Health Survey (SF-36) (89,90), and items from the Symptom Checklist 92 (SCL-92) (91), or self-

constructed ad hoc items. The questionnaire also includes an item regarding the patient's own perceived need of contact to the outpatient clinic to ensure that patients can get an appointment if they want to, regardless of their response to the other items in the questionnaire.

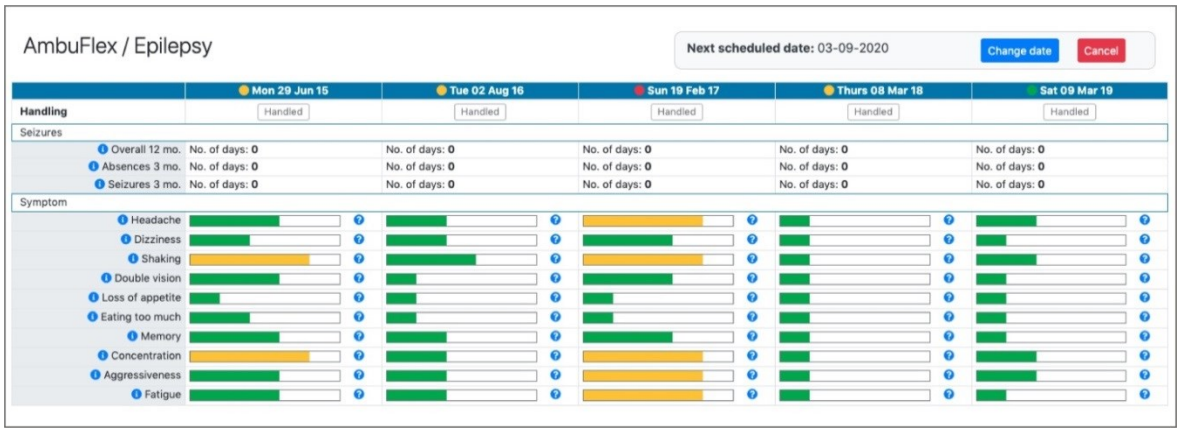
The PRO-based algorithm

A clinical expert group has marked all the item response categories in the epilepsy questionnaire with either a green, yellow, or red color based on what the clinicians considered of clinically important to react to, in order to identify patients who need clinical attention (4). The algorithm is based on a "red flag" approach; if a red color is given to a response category, it indicates a need of clinical attention. For example, planning pregnancy, seizure impairment, suicidal thoughts, or if the patient wishes contact. The yellow color is given if the item response category is of minor concern, for example, reporting seizures, symptoms related to side-effects, or mental problems. The green color is only given to item response categories which indicate no problems at all. Thus, the overall questionnaire response results in a color that indicates how the clinicians should review the responses. A green color indicates no need of clinical attention; a yellow color indicates possible need of attention; and a red color indicates need of attention (4). A gray color is given to non-responders after three reminders. All green responses are managed automatically by the AmbuFlex web system, and a new questionnaire is sent to the patient at the pre-defined fixed interval, e.g. after 6 or 12 months. As a safety precaution, a patient with only green responses will turn yellow after 3 years.

Review of responses

The questionnaire responses are presented in a graphic PRO overview (Figure 4), which is available for the clinicians via the Electronic Health Record system together with other relevant data in the patient's record (e.g. blood test results, medication, etc.) (4). All the red, yellow, and gray questionnaire responses are shown to the clinicians on an alert list. For red responses, the clinicians contact the patient as quickly as possible either by telephone or by scheduling an in-clinic appointment. For yellow responses, the clinicians evaluate the PRO data together with other available data from the medical record. They only contact the patient if they find it necessary. Non-responders are contacted by clinicians based on a local procedure in each department. Table 4 presents an overview of PRO responses and distribution of the PRO-based algorithm in the Central Denmark Region during the period from March 2012 to January 2020. In all, 54% of the PRO responses in this period have resulted in no further contact to the outpatient clinic.

Figure 4 The PRO overview shown to the clinicians in the Electronic Health Record (4)



The PRO-based algorithm is shown by the color dots in the upper row (green: no need of attention; yellow: possible need of attention; and red: definite need of attention). The color of the bars is identical with the PRO-based algorithm for each of the questions shown. The length of the bars illustrates the degree of symptom severity. Note: the text in the figure has been translated from Danish.

Table 4 Distribution of the PRO-based algorithm in outpatients with epilepsy in the Central Denmark Region from 2012 to 2020, *n*(%)

	Aarhus University Hospital	**Regional Hospital Holstebro**	**Regional Hospital Viborg**	**Total**
PRO responses	14921 (100)	2770 (100)	967 (100)	18658 (100)
PRO-based algorithm				
Green	2060 (14)	609 (22)	172 (18)	2841(15)
Yellow	9383 (63)	1658 (60)	661 (68)	11702 (63)
Red	3478 (23)	503 (18)	134 (14)	4115 (22)
No contact	7468 (50)	1942 (70)	666 (69)	10076 (54)
Contact	7448 (50)	792 (29)	289 (30)	8529 (46)
Pending*	5 (0.03)	36 (1)	12 (1)	53 (0.3)

*Await clinical assessment

SYNTHESIS

PRO measures are increasingly being used at the individual patient level in the healthcare system. AmbuFlex is a model of remote PRO-based follow-up, and the model includes different elements to support efficient use of PRO data in remote patient management. However, the evidence related to use of PRO measures in remote PRO-based follow-up is scant. Therefore, we set out to investigate different aspects related to the pathway of using PRO-based follow-up in outpatients with epilepsy. Research is of major importance to all

these aspects in order to ensure development and implementation of high-quality PRO-based follow-up solutions. We decided to focus on three aspects.

Referral to PRO-based follow-up

Since 2012, approximately 50% of the outpatient epilepsy population in the Central Denmark Region has been referred to PRO-based follow-up. Patients are individually referred to PRO-based follow-up based on the patient's preferences and clinical profile. Research on this topic is scant, and I have not been able to identify studies that have investigated the characteristics of patients who attend PRO-based follow-up. However, the characteristics of questionnaire non-responders could be relevant to consider, and studies have reported non-response to be associated with, for example, male sex, younger age, lower socioeconomic status, living alone, and poorer health status (92-98). Research on this topic is relevant to healthcare planning when using PRO measures as a strategy in remote outpatient follow-up.

Reliability of the epilepsy questionnaire and the PRO-based algorithm

The epilepsy questionnaire and the PRO-based algorithm have been developed and tested in close cooperation with patients and clinicians and have been used in routine clinical practice since 2012. During the questionnaire development process, content and face validity were the primary focus, though validity and reliability have not yet been documented (4). However, psychometric properties, such as reliability needs to be further evaluated to ensure the consistency of both the included items and the PRO-based algorithm. Moreover, the reliability of the WHO-5 scale has been evaluated in terms of internal consistency in other studies (99-103); however, test-retest reliability and measurement error of the WHO-5 scale have not been evaluated sufficiently.

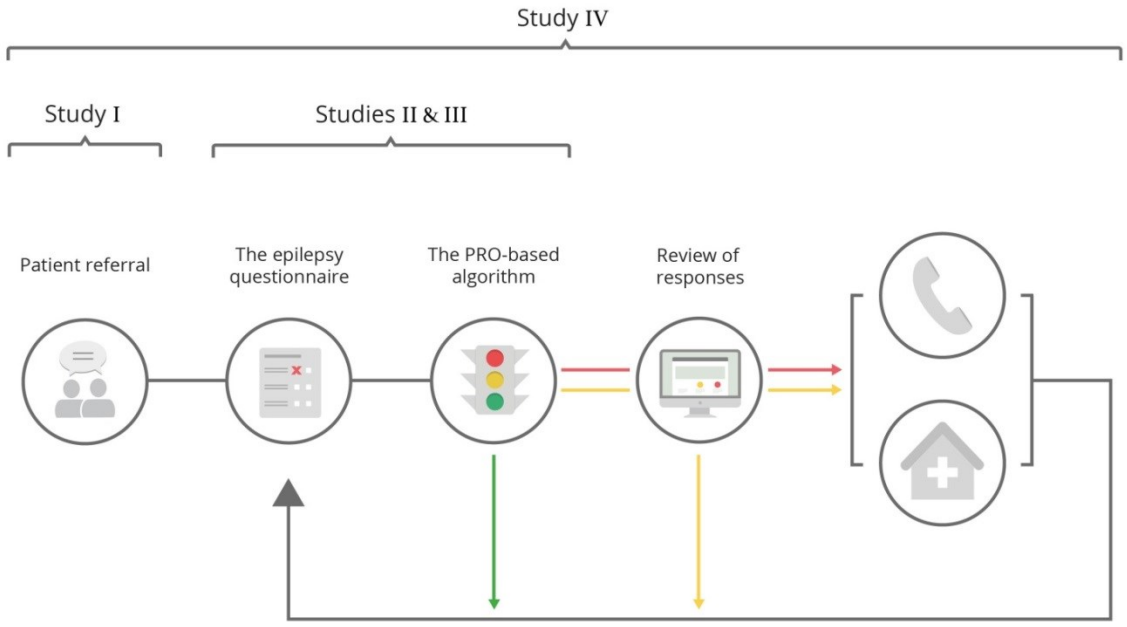
The effect of patient-initiated PRO-based follow-up

In 2015, PRO-based follow-up using fixed questionnaire intervals was standard follow-up of approximately 2500 outpatients with epilepsy at Aarhus University Hospital. However, fixed-interval PRO-based follow-up may not be an adequate solution if patients have fluctuating disease activity. Therefore, we found it reasonable to consider the benefits of using patient-initiated PRO-based follow-up using non-scheduled patient-initiated questionnaire intervals in a cohort of patients attending fixed-interval PRO-based follow-up. In so doing, we focused not only on reducing the use of healthcare services but also on providing patient-centered care in which the patients are empowered to react promptly to symptom impairment and decide the need for outpatient contacts. To provide knowledge to guide the use of future remote PRO-based interventions, possible benefits and drawbacks of the effect of patient-initiated PRO-based follow-up need to be evaluated.

AIMS

This dissertation set out to provide further insight into different aspects of the use of remote PRO-based follow-up in outpatients with epilepsy. Figure 5 illustrates the aims according to the PRO-based follow-up model.

Figure 5 Aims of the studies according to the PRO-based follow-up model



Study I aimed to identify sociodemographic, personal, and disease-related factors associated with referral to PRO-based follow-up (Paper I).

Study II aimed to evaluate the test-retest reliability of the PRO-based algorithm used for clinical decision support in epilepsy outpatient follow-up and to analyze whether different methods of administration (web, paper, or a mixture of the two modalities) influenced the results. Moreover, the study aimed to evaluate the test-retest reliability of the single items included in the epilepsy questionnaire (Paper II).

Study III aimed to evaluate the test-retest reliability and measurement error of the Danish version of the WHO-Five Well-Being Index in outpatients with epilepsy and to evaluate to what extent different methods of administration (web, paper, or a mixture of the two modalities) influenced the results (Paper III).

Study IV aimed to evaluate the effects of PRO-based patient-initiated follow-up in outpatients with epilepsy. The study aimed to compare use of healthcare resources, quality of care, and the patient perspective in two outpatient follow-up activities: patient-initiated PRO-based follow-up and fixed-interval PRO-based follow-up (Papers IV & V).

Hypotheses

In Study I, we hypothesized that a low level of health literacy, self-efficacy, patient activation, general health, well-being, education, household income and higher age, solo living, passive labor market participation, and a high level of co-morbidity were associated with lower probability of referral to PRO-based follow-up.

In Study IV, we hypothesized that the use of healthcare resources would be lower, quality of care at least as good and patient self-management and patient satisfaction would be improved among patients in the patient-initiated PRO-based intervention arm compared with those in the fixed-interval PRO-based control arm.

We did not construct hypotheses in Studies II & III.

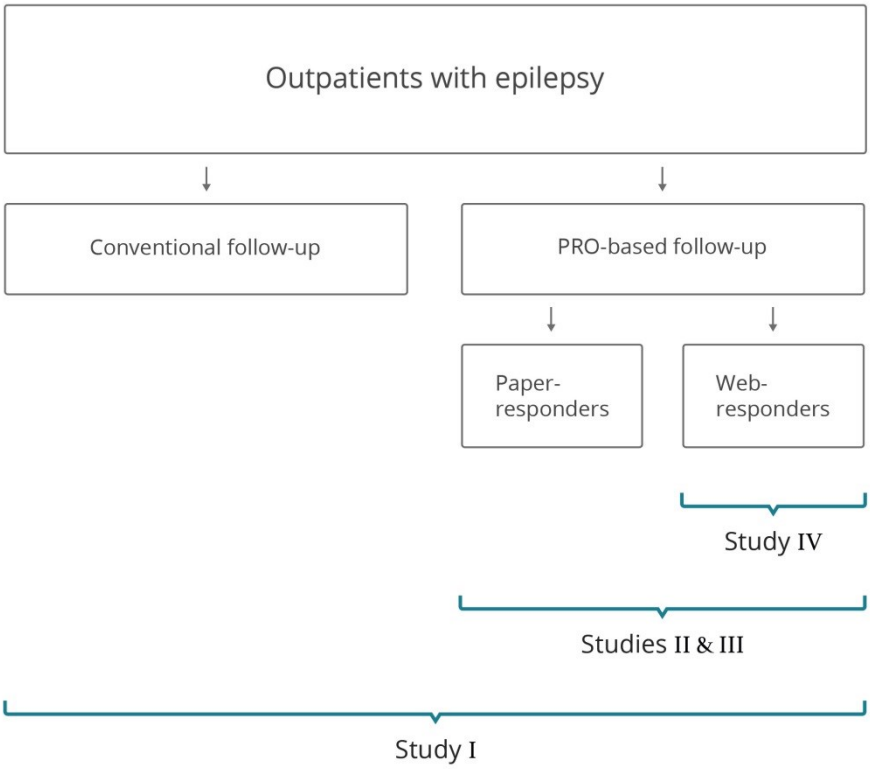
METHODS

This dissertation is based on four studies applying three different study designs and conducted in three different study populations. In this section, materials and methods will be presented separately for Study I and Study IV, and combined for Studies II & III. For further details, see the method sections of the appended papers. Table 5 presents an overview of study design methodology, and Figure 6 presents an overview of the study populations in the four studies.

Table 5 Overview of study design methodology in the four studies

	Study I	**Study II**	**Study III**	**Study IV**
Topic	Determinants of referral to PRO-based follow-up	Test-retest reliability of a PRO-based algorithm used for clinical decision support	Test-retest reliability and measurement error of the Danish WHO-Five Well-Being Index	The effectiveness of patient-initiated PRO-based follow-up
Design	Cohort study	Test-retest study	Test-retest study	Randomized controlled trial
Patients	*N* = 802	*N* = 554	*N* = 554	*N* = 593
Data sources	National and regional registers and questionnaires	Questionnaires	Questionnaires	Regional register and questionnaires
Independent variables (exposures)	Age, gender, cohabitation status, education, household income, labor marked affiliation, co-morbidity, psychiatric disease, well-being, general health, health literacy, self-efficacy, and patient activation.	N/A	N/A	Patient-initiated PRO-based follow-up versus fixed-interval PRO-based follow-up (standard care)
Dependent variables (outcomes)	Referral to PRO-based follow-up	N/A	N/A	Number of outpatient contacts, hospital admissions, emergency room visits, mortality, well-being, general health, number of seizures, side-effects, health literacy, self-efficacy, patient activation, confidence, safety, and satisfaction
Data analysis	Time to event analysis using the pseudo-value regression approach	Proportion and unweighted and weighted kappa	Intraclass correlation coefficient Weighted kappa Standard error of the measurement	Group comparison using linear regression models
Paper	I	II	III	IV & V

Figure 6 Overview of study populations in the four studies



LITERATURE SEARCH

A systematic literature search was conducted according to specific themes and papers (Appendix 1). The search strategy was developed in cooperation with a librarian from the Hospital Library in the Central Denmark Region. The search is based on a thematic block search and conducted from 2015 onwards, mainly in the Pubmed database. Moreover, references from relevant papers were reviewed to identify additional relevant literature.

ETHICS

This dissertation adheres to Danish ethical research standards and the Helsinki Declaration. In 2015, the PhD project was authorized and approved by the Danish Data Protection Agency (reference number 1-16-02-691-14). Moreover, the Ethics Committee of the Central Denmark Region was contacted and responded that according to Danish law the study did not require approval from the Committee, as human biological material was not included (§ 14) (104). In Study I, results were handled anonymous in Statistics Denmark (DST number: 706981). Study IV was registered with ClinicalTrials.gov: number NCT02673580. A license agreement was obtained to use the Health Literacy Questionnaire

in Study I and Study IV. In accordance with guidelines from the Danish Data Protection Agency, consent to participate in the studies was informed, specific, voluntary, and explicit (105). In all four studies, patient information was given to the patient in a written form together with a research questionnaire. The patient information also stated that participation consent could be withdrawn at any time. The data were handled and stored with confidentiality.

STATISTICAL PROGRAM

All statistical analyses were carried out using the statistical package STATA version 15 (Stata Corporation, College Station, Texas, USA).

STUDY I: DETERMINANTS OF REFERRAL TO PRO-BASED FOLLOW-UP

Study design and population

Study I was designed as a prospective cohort study in epilepsy outpatients at the Department of Neurology at Aarhus University Hospital, Denmark. Using the Hospital Business Intelligence (BI) Register in the Central Denmark Region (106), every second week during the period from May 2016 to May 2018 we identified all patients aged ≥15 years who were visiting the department for the first time and were diagnosed with epilepsy or suspicion of epilepsy according to the 10th version of the International Classification of Diseases (ICD-10): DG40–409, DZ033A, DR568E, and DR568.

Questionnaire

A questionnaire was sent to all eligible patients approximately 2 weeks after their first visit at the department. The questionnaire covered a range of self-reported measures including subscales 4, 6, and 9 from the Health Literacy Questionnaire (HLQ) (107,108), the General Self-Efficacy Scale (GSES) (109-111), the WHO-Five Well-Being Index (WHO-5) (87,88), a single item from the Short Form 36 Health Survey (SF-36) (89,90), and two items modified from the Patient Activation Measure 13 (PAM-13) (112). Patients completed either a paper or web version of the questionnaire. One reminder was sent after 21 days to non-responders.

Registers

Supplementary to the questionnaire, we had access to the BI Register and a variety of established Danish registers via Statistic Denmark (113). Information from the registers can be combined at the individual level by means of a unique personal identification (CPR) number given to all Danish citizens at birth and to residents upon immigration

(114,115). In January 2019, questionnaire data and data from the BI Register were linked with register data in Statistics Denmark.

The Hospital Business Intelligence (BI) Register

Since 2011, the Hospital BI Register has incorporated information about hospital-based activities including hospital admissions, emergency room visits, and outpatient visits from the public hospitals in the Central Denmark Region (106). Both primary and secondary diagnoses are registered along with information on procedures and treatments, and data are updated on a daily basis. We used the BI Register to identify eligible study participants. Gender and date of birth was provided by information in the CPR number, and we retrieved information on the date of exclusion of outpatient care, emigration, and death from the register.

The Danish Civil Registration System

The Danish Civil Registration System contains information about vital status and addresses of all Danish citizens and is updated daily (114). The register was used for two different purposes. We retrieved information on the included participants' addresses and vital status prior to sending a questionnaire to a participant via the AmbuFlex system. Secondly, we retrieved information on cohabitation status the year before inclusion in the study from Statistics Denmark.

The Danish Education Registers

The Danish Education Registers contain information on individuals' highest completed education for 96.4% of the Danish population between 15 and 69 years of age (116). The Danish educational institutions provide information about individual-level status on enrollment, exams, and completed levels of education on a yearly basis. We retrieved information on the participants' highest level of completed education the year before inclusion in the study.

The Danish Registers on Income and Transfer Payments

The Danish Registers on Personal Income and Transfer Payments include many variables that provide an overview of income composition for the entire Danish population (117). We retrieved information on yearly household income and equivalised disposable household income for the participants' residence the year before inclusion in the study. Equivalised household income takes into account differences in composition of the household.

The Danish Register for Evaluation and Marginalization (DREAM)

The DREAM contains information on a persons' attachment to the labor market and many temporary and permanent social benefits and is updated on a weekly basis (118). We retrieved information on labor market participation for the participants within a 52-week period from before the date of inclusion in the study.

The Danish National Patient Register (DNPR)

The DNPR provides information on all diagnoses registered in relation to hospital admissions (from 1977), emergency room visits and outpatient visits (from 1995) in private and public hospitals in Denmark (119). We retrieved information on participants' co-morbidity within 10 years before enrollment and psychiatric diseases within 2 years before enrollment in the study.

Determinant variables

Determinants or independent variables included both questionnaire- and register-based data. Table 6 shows an overview of the determinant variables and the data sources used in the study (120).

Table 6 Overview of determinant variables and register and questionnaire data sources (120)

Determinant	Data source
Age Gender	The Hospital Business Intelligence (BI) Register in Central Denmark Region
Cohabitation status	The Danish Civil Registration System
Education	The Danish Education Registers
Household income	Danish Registers on Income and Transfer Payments
Labor market affiliation	The Danish Register for Evaluation and Marginalization (DREAM)
Co-morbidity Psychiatric disease	The Danish National Patient Register (DNPR)
Well-being	WHO-Five Well-Being Index (WHO-5)
General health	Short Form 36 Health Survey (SF-36)
Health literacy	Health Literacy Questionnaire (HLQ)
Self-efficacy	General Self-Efficacy Scale (GSES)
Patient activation	Patient Activation Measure 13 (PAM-13)

Register-based data

Age was measured at the date of inclusion in the study and categorized into five age groups. **Cohabitation status** was categorized into: "Living with a partner/family" or "Living alone". **Education** was categorized into three groups: low (< 10 years), medium (10–12 years), or high (> 12 years) educational level based on the International Standard Classification of Education (121). **Household income** was categorized into three groups according to tertiles (33.3rd and 66.6th percentile): low, medium, or high income level. **Labor market affiliation** was categorized into four groups: self-supporting, normal retirement, receiving temporary social benefits, and receiving permanent social benefits. **Co-morbidity** was measured using the Charlson Comorbidity Index that categorized the participants into three levels of co-morbidity: 0 (Low), 1–2 (Medium), and >2 (High) (122). **Psychiatric diseases** were dichotomized into two groups: present or not.

Questionnaire-based data

Health literacy was measured using the HLQ average scores across all items from the subscales 4: Social support for health (five items); 6: Ability to actively engage with healthcare providers (five items); and 9: Understanding health information well enough to know what to do (five items). Higher subscale scores indicate better degrees of health literacy (107,108). **Self-efficacy** was measured using the GSES average score across all 10 items in the scale. A higher score indicates a better degree of self-efficacy (109-111). **Well-being** was measured using the WHO-5 average score across all five items in the scale. A higher WHO-5 score indicates a better degree of well-being (87,88). All scale scores were analyzed in their continuous form and dichotomized. For example, a WHO-5 percentage score below 50 indicates increased risk of depression (123); therefore, we dichotomized the WHO-5 score at 50, as the potential association would be easier to interpret. However, due to risk of loss of information, we also decided to analyze the scores continuous. Further details can be found in Paper I.

General health was measured using a single item from SF-36: "In general, would you say your health is: excellent, very good, good, fair, or poor". The variable was categorized into three groups: "excellent/very good", "good", and "fair/poor" (89,90). **Patient activation** was measured using two single items modified from the PAM-13: "I am confident that I can tell when I need to get outpatient care" and "I am confident I can figure out solutions when new situations or problems arise with my health condition" with the response categories: "disagree strongly", "disagree", "agree", and "agree strongly" (112). The variables were categorized into "disagree strongly/ disagree" and "agree/agree strongly".

Outcome variable

The outcome or event of interest was referral to PRO-based follow-up after the participants' first visit at the Department of Neurology at Aarhus University Hospital. The proportion of patients referred to PRO-based follow-up was evaluated within 6, 12, and 18 months after their first visit at the department. Patients were censored after 18 months follow-up or in January 2019, or if the patients had no further need of outpatient care, emigrated, or died. These data were extracted from the regional Hospital BI Register. The date of referral to PRO-based follow-up was extracted from the AmbuFlex database (4,17).

Statistical analysis

Power calculation

We assumed that approximately 50% of the patients were referred to PRO-based follow-up. To detect a 15% difference where 55% patients with "excellent/very good" general health and 40% patients with "fair/poor" general health were referred to PRO-based follow-up, the power was estimated to be 92% with a sample size of 500 patients and a *p*-value of 0.05. The risk of missing a real effect is then 8%. The power was estimated to be 74% if the total sample size number was reduced to 300.

Pseudo-value regression

Associations between determinants and referral to PRO-based follow-up were analyzed using the pseudo-value regression approach estimating the cumulative risk ratio (RR) of PRO-based follow-up at three time points: 6, 12, and 18 months after the participants' first visit at the Department of Neurology at Aarhus University Hospital. The pseudo-value regression approach involves a new set of observations/pseudo-values that are generated and used in a generalized linear regression model (124,125). Moreover, the model can take competing risks into account. In our study, competing risk factors included death, emigration, and end of outpatient follow-up if these events took place before referral to PRO-based follow-up. For all analyses, a p-value < 0.05 was considered statistically significant, and 95% confidence intervals (CI) were reported. Confounder variables included age, gender, cohabitation status, education level, and co-morbidity, which were identified a priori based on previous studies about factors associated with questionnaire non-response (92-98).

Multiple imputation

We used multiple imputation (MI) to handle the missing data problem in both the questionnaire and register data (126,127). MI is a statistical approach that can be used if data are missing at random (MAR) (126). In MAR, differences in observed data may be used to explain any systematic differences between observed and missing data (127). For example, missing health literacy measures may be lower than measured health literacy, but only because lower educated people may be more likely to have missing health literacy measures. Based on a MI model of relevant variables measured in the study population and under the assumption that data were MAR, we created 100 complete datasets. Two other MI models were created by modifying the observed variables to evaluate the robustness of the first model. The three models are presented in Appendix 1 of Paper I. It is not possible to distinguish between MAR and missing not at random (MNAR); thus, biases caused by data that are MNAR can be addressed only by sensitivity analyses (127). We conducted sensitivity analyses in which we assumed that health literacy data were MNAR.

STUDY II AND STUDY III: TEST-RETEST RELIABILITY OF THE PRO-BASED ALGORITHM AND WHO-5

Study design and population

Studies II & III were designed as test-retest reliability studies among outpatients with epilepsy from three neurological departments in the Central Denmark Region. The departments are located at Aarhus University Hospital, Regional Hospital Holstebro, and Regional Hospital Viborg. Patients were included if they were ≥ 15 years and were attending PRO-based follow-up in the period from August 2016 to April 2017. The participants responded to the disease-specific epilepsy questionnaire at two time points: at test 1 and test 2.

Test 1: Firstly, the participants responded to a scheduled questionnaire from the outpatient clinic. Participants completed either a web or paper version of the questionnaire based on their preferred method of administration. Three reminders were sent to non-responders.

Test 2: Secondly, the participants responded to the same questionnaire approximately 2 weeks after their first response. Based on their response method in test 1, the participants were randomly assigned to fill in either a web or paper version of the second questionnaire. No reminders were sent to non-responders.

In summary, based on the methods of administration of test 1 and test 2, participants were divided into four test-retest groups: web-web, paper-paper, web-paper, and paper-web.

Development of the epilepsy questionnaire and the PRO-based algorithm

In 2011, a disease-specific questionnaire was developed that aimed to identify epilepsy patients' health problems in order to support clinical decision-making in remote outpatient follow-up (Appendix 2). The epilepsy questionnaire was developed and tested in close cooperation with patients and clinicians. They are experts and know best which aspects to include in the questionnaire in order to using it for clinical decision-making. It was pivotal to develop a questionnaire that are clinical relevant to both patients and clinicians (4). The primary focuses during the development process was content and face validity (22,128). The development process was pragmatic and based on consensus decision-making involving several face-to-face meetings. Overall, the process involved three iterative phases: 1. defining aim, content and construction of the questionnaire, 2. pilot-testing, and 3. defining the PRO-based algorithm.

First, a working group including clinical experts of epilepsy and PRO experts defined the aim and the target population of the questionnaire. Thereafter, clinical experts of epilepsy and PRO experts contributed with suggestions regarding the content and construction of

the questionnaire. A systematic literature search identified validated PRO measures that could be used to measure some of the included constructs, e.g. well-being was measured using the WHO-5 (87,88), general health was measured using items from SF-36 (89,90), and symptoms were measured using items from the SCL-92 (91). If established PRO measures were not identified in the literature, self-constructed ad hoc items were developed. Clinicians provided inputs to item content, and PRO experts provided inputs to item formulation and scoring using established item response categories if appropriate.

Second, the first draft of the questionnaire was pilot-tested by 20 epilepsy outpatients using cognitive semi-structural interviews (129). The purpose was to test understanding of items, relevance of items, and lack of relevant themes (22). The pilot-test did not identify any major problems as the patients found the content of the questionnaire relevant and without critical problems regarding understanding (4).

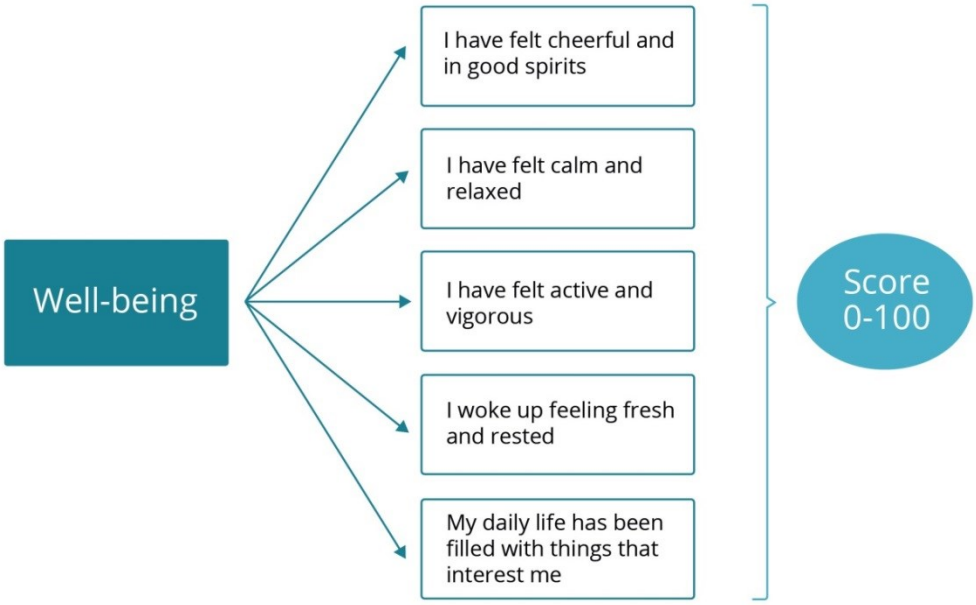
Finally, based on the first draft of the epilepsy questionnaire, a clinical expert group developed a PRO-based algorithm, whereby they classified all the item response categories into a green, yellow, or red color based on what they found most important to react to in order to identify patients who need clinical attention (4). The PRO-based algorithm is based on a "red flag" approach in which specific answers of clinical importance are given a red color indicating a need for clinical attention, e.g. suicidal thoughts, seizure impairment, and pregnancy.

Both the epilepsy questionnaire and PRO-based algorithm were evaluated and revised yearly at consensus meetings, until 2016.

The WHO-Five Well-Being Index

The construct "well-being" was measured using the generic questionnaire WHO-Five Well-Being Index (WHO-5) (87,88). The WHO-5 has a single factor structure and consists of five items (Figure 7). The five items have six ordinal response categories ranging from 0 "At no time" to 5 "All of the time". The total raw score is estimated by summing item scores (range from 0 to 25) and a total percentage score is estimated by multiplying the raw score by four (range from 0 to 100). Lower scores indicate impaired emotional well-being and a raw score < 13, or percentage score < 50, or a score of 0 or 1 in any of the items indicate impaired well-being and risk of depression (123). The unidimensional structure of the WHO-5 has been confirmed by the item response theory Rasch model in several studies (88). Moreover, other psychometric properties, such as responsiveness in clinical trials, predictive validity, and criterion validity of the WHO-5 as a screening tool for depression have been evaluated to be adequate (88). The WHO-5 has been translated into 30 language versions and has been applied across many different patient populations and clinical settings (88).

Figure 7 The unidimensional structure of the WHO-Five Well-Being Index (WHO-5)



Statistical analysis

Test-retest reliability was assessed by comparing the two questionnaire responses (test 1 and test 2). The test-retest interval was estimated by calculating the number of days between the two responses. Test-retest reliability was also assessed according to the different methods of administration in both Study II and Study III. According to the COSMIN checklist for studies assessing validity and reliability, a sample size of at least 50 participants was considered to be sufficient (24,130). Missing items were omitted from the analyses, and the WHO-5 score was not calculated if items were missing.

Kappa

In Study II, we used unweighted kappa (in nominal data) and weighted kappa with squared weights (in ordinal data) to assess test-retest reliability (22,131). We assessed both test-retest reliability according to the PRO-based algorithm (green, yellow, or red colors) and the single items included in the epilepsy questionnaire. The following recommendation from Landis et al. was used in the interpretation of the kappa coefficients: <0.2 (slight), 0.21–0.4 (fair), 0.41–0.6 (moderate), 0.61–0.8 (substantial), and 0.81–1.0 (almost perfect) (132). In Study III, we used weighted kappa statistics with squared weights to assess test-retest reliability of the five single WHO-5 items.

Intraclass correlation coefficient (ICC)

In Study III, we used intraclass correlation coefficient (ICC) agreement model 2.1 (133) to assess test-retest reliability of the WHO-5 scale. At group level, an ICC of 0.70 is considered acceptable, though at the patient level an ICC of 0.90 is recommended (22).

Standard error of the measurement

In Study III, we used standard error of the measurement to assess measurement error. First, we illustrated the differences between test 1 and test 2 using the Bland-Altman plot in which the differences were plotted against the means of the two test-retest measurements (22). Second, we estimated the standard error of the measurement which reflects the intra-individual variation (134). Based on the standard error of the measurement, we estimated the minimal detectable change (MDC) (134). The MDC points to the smallest within-person change that can be explained as a real individual change over the measurement error (134); hence, a change in scores smaller than the MDC can be ascribed to measurement error and may not be a real change.

Sensitivity analyses and attrition

The time interval between test 1 and test 2 is of importance due to either recall or a real change in the patients' health status. The patients were not asked whether their health status had changed between the two time points. However, we performed sensitivity analyses in both Study II and Study III to assess whether the length of the time interval affected the results. We analyzed differences between responders and non-responders of the second questionnaire based on data from the first questionnaire. We used a chi-squared test for categorical variables and the Kruskal-Wallis test for continuous variables.

STUDY IV: THE EFFECTIVENESS OF PATIENT-INITIATED PRO-BASED FOLLOW-UP

Study design and population

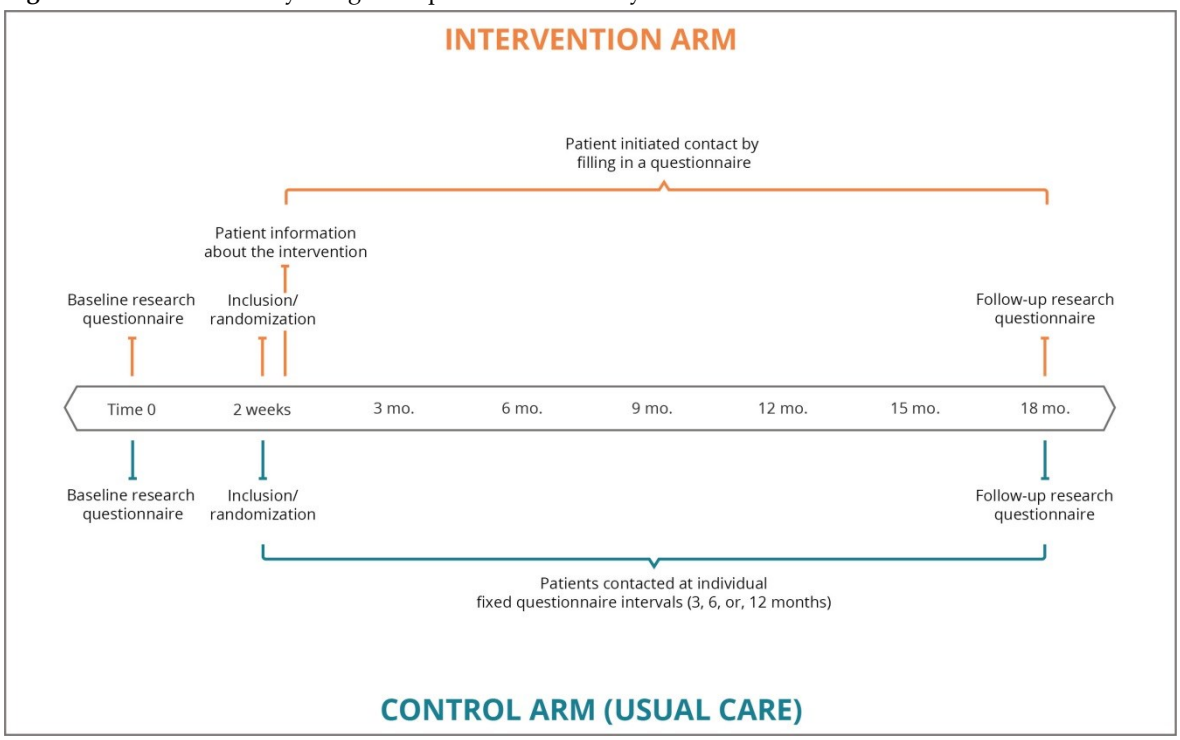
Study IV was designed as a pragmatic two-armed parallel randomized controlled study with an intervention arm and a control arm. We followed the Consolidated Standard of Reporting Trial (CONSORT) checklist in reporting parallel arm randomized trials (135) and the CONSORT PRO extension (136). The trial included outpatients with epilepsy at the Department of Neurology at Aarhus University Hospital, Denmark. Patients were included if they were attending fixed-interval PRO-based follow-up, 15 years of age or older, and web-responders.

Recruitment and procedures

From January 2016 to July 2016, all patients attending fixed-interval PRO-based follow-up received a research questionnaire combined with their usual epilepsy questionnaire from the department. Patients filled in the questionnaire either in a paper- or a web-based form. Approximately 2 weeks after the questionnaire response, the patients who were web-responders were randomized to the intervention arm or control arm. Patients in the

intervention arm subsequently received detailed written information about the intervention and guidance to use the intervention website 'My Epilepsy' (Appendix 3). They were also informed that they could contact the study coordinator if they did not want to participate and wanted to continue fixed-interval PRO-based follow-up. Patients in the control arm continued usual care and no changes were implemented. Control arm patients continued to receive the epilepsy questionnaire at fixed intervals during follow-up. These fixed intervals were not the same for all patients, for example, some patients filled in the questionnaire every 3rd month, some every 6th month, but the majority of patients in the control arm filled in a scheduled epilepsy questionnaire once a year. Both arms received a follow-up research questionnaire 18 months after randomization. Figure 8 presents an overview of the study design and procedures.

Figure 8 Overview of study design and procedures in Study IV



Randomization and blinding

Eligible patients were pre-randomized in a ratio of 0.55:0.45 to either the intervention arm: patient-initiated PRO-based follow-up (open access telePRO); or the control arm: fixed-interval PRO-based follow-up (standard telePRO). We used a pre-randomization design since all patients were already attending fixed-interval PRO-based follow-up (137). Patients in the intervention arm were informed about the allocation following randomization and no information about allocation was provided to patients in the control arm who continued the standard follow-up activity with no changes. We used simple and computer-generated randomization using the AmbuFlex web system (17).

Blinding of the patients or those who provided the intervention was not possible due to the nature of the intervention.

Intervention

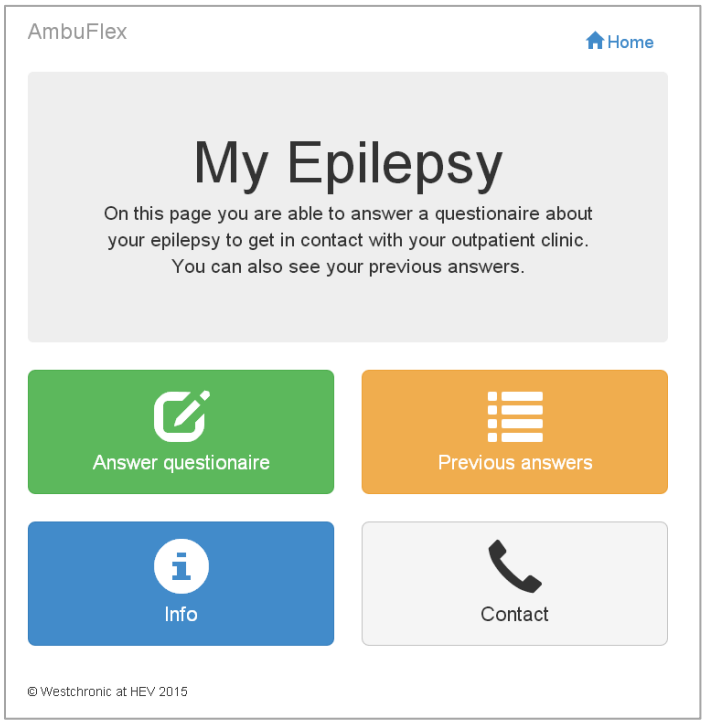
In the intervention arm, fixed-interval questionnaires were replaced by a patient-initiated follow-up method. This group of patients evaluated their own need of clinical attention if they had symptoms or other concerns, and initiated contact with the outpatient clinic by filling in a disease-specific questionnaire. A website called 'My Epilepsy' was developed for this purpose.

The development of the 'My Epilepsy' website followed three steps:

1. Initial draft of the content and design of the prototype website by including outpatients with epilepsy, clinical experts, researchers, and software developers.
2. Usability testing using cognitive interviewing techniques to collect feedback from outpatients with epilepsy.
3. Final revision of the website based on the results from the patient interviews.

The final version contains four core elements (Figure 9) that gave patients access to: 1. fill in the epilepsy questionnaire; 2. an overview of previously questionnaire responses; 3. information about the website and the questionnaire; and 4. telephone number to the epilepsy outpatient clinic (138). Further details about the intervention and the website can be found in Paper IV and Paper V.

Figure 9 The patient-initiated/open access website 'My Epilepsy' (138)



Patients accessed the website via a secure login procedure to a Danish National eHealth Portal (https:\\Sundhed.dk). If the patient filled in a questionnaire, the response was fed back to clinicians who accessed the response via an alert list in the AmbuFlex web system. All patient-initiated responses were given a red algorithm color. A clinician reviewed the patient's response and called the patient as soon as possible, and in some cases an in-clinic visit was arranged. As a safety precaution, reminders to fill in the questionnaire were sent to patients who did not fill in the questionnaire for a prolonged period; for example, a reminder was sent after 12 months if a patient was referred with a 6-month questionnaire interval to PRO-based follow-up prior to randomization.

Outcome measures

The primary outcome measure was number of outpatient contacts related to epilepsy in an 18-month time span from baseline to follow-up. Outpatient contacts included both telephone consultations and outpatient visits, which were evaluated separately. Data were collected from the Hospital BI Register in Central Denmark Region (106). Data regarding hospital admissions, emergency room visits, and mortality were also collected from the BI Register.

Patients were also assessed at baseline and at 18-month follow-up by completion of a questionnaire that included several secondary outcomes. These were selected patient health outcomes such as well-being evaluated by the WHO-5 (87,88), general health evaluated by one item from SF-36 (89,90), number of seizures and treatment side effects evaluated by single items from the epilepsy questionnaire used in AmbuFlex (Appendix 2). The patient perspective included measures related to self-management, such as health literacy evaluated by HLQ (subscale 4, 6, and 9) (107,108), self-efficacy evaluated by the GSES (109-111), and patient activation evaluated by two items modified from the PAM-13 (112). Three single item measures related to confidence, safety, and satisfaction modified from a satisfaction questionnaire from the Danish Cancer Society were used in health service evaluation (139). An overview of the secondary self-reported outcomes, data measures, and scoring can be found in Table 7.

Other measurements such as education level, cohabitation status, and duration of epilepsy were collected from the baseline research questionnaire.

Table 7 Secondary self-reported outcomes, data measures, and scoring

Secondary outcomes	Data measures	Items	Scoring
Well-being	WHO-Five Well-Being Index (WHO-5)	5 items	0 (worst) – 100
General health	The Short Form 36 Health Survey (SF-36)	1 item	1 (best) – 5
Number of seizures	The epilepsy questionnaire, AmbuFlex	1 item	NA
Treatment side effects	The epilepsy questionnaire, AmbuFlex	1 item	1 (best) – 4
Health literacy	The Health Literacy Questionnaire (HLQ): Subscale 4: Social support for health	5 items	1 (worst) – 4
	Subscale 6: Ability to actively engage with health care providers	5 items	1 (worst) – 5
	Subscale 9: Understanding health information well enough to know what to do	5 items	1 (worst) – 5
Self-efficacy	General Self-Efficacy Scale (GSES)	10 items	10 (worst) – 40
Patient activation	Patient Activation Measure (PAM): I am confident that I can tell when I need to get outpatient care	1 item	1 (worst) – 4
	I am confident I can figure out solutions when new situations or problems arise with my health condition	1 item	1 (worst) – 4
Confidence	Satisfaction questionnaire, Danish Cancer Society	1 item	1 (best) – 4
Safety	Satisfaction questionnaire, Danish Cancer Society	1 item	1 (best) – 4
Satisfaction	Satisfaction questionnaire, Danish Cancer Society	1 item	1 (best) – 4

References (87-90,107-112,139)

Statistical analysis

Sample size

A sample size calculation was performed before initiation of the trial. The calculation was based on a two-sided statistical test with 90% power and a 5% level of significance, a difference between the two arms of 1 outpatient contact, and a standard deviation (SD) of 3.41 in the intervention arm and a SD of 2.38 in the control arm. The SDs were based on data from a RCT study that investigated the effect of an open access intervention (140). The calculation led to an estimated sample size of 386 patients in total.

Linear regression

The main analysis was performed according to the intention-to-treat (ITT) approach and was supported by per protocol analysis. The effect was analyzed using a linear regression model for both primary and secondary outcomes. The bootstrap method was used to estimate 95% confidence intervals if the normality distributions were skewed (141). A simple linear regression model was used to estimate ITT between-arm differences of primary outcome, and a multiple linear regression model was used to estimate ITT between-arm differences of secondary outcomes by calculating differences at 18-month follow-up adjusted for the baseline value. Furthermore, a multiple linear regression model

was used to estimate per protocol between-arm differences for both primary and secondary outcomes. These models included adjustment for age, gender, education level, cohabitation status, duration of epilepsy, and number of seizures during last year. We also performed explorative ITT analyses by stratifying age, gender, and high/low health literacy (subscale 4) and sensitivity analyses to investigate the robustness of the ITT results related to self-reported well-being (WHO-5).

Baseline characteristics and attrition

We analyzed differences in baseline characteristics, differences between paper and web responders, and differences between responders and non-responders of the follow-up research questionnaire. We used a chi-squared test for categorical variables and the Wilcoxon Mann-Whitney test or unpaired *t*-test for continuous variables.

RESULTS

This section presents the main results of the four studies. Results are presented separately for Study I and Study IV, and combined for Studies II & III. Additional results and more detailed presentations are available in the appended papers and supplemental materials contained in the papers. Table 8 presents an overview of the study population characteristics in the four studies.

Table 8 Overview of study population characteristics in the four studies

	Study I [a]	**Studies II & III** [b]	**Study IV** [c]
Population	*N* = 802	*N* = 554	*N* = 593
Gender, male *n* (%)	415 (52)	286 (52)	297 (50)
Age, years, mean (SD)	49.3 (21.9)	55.0 (16.6)	46.6 (17.2)
Outpatient clinic, *n* (%)			
Aarhus	802 (100)	409 (74)	593 (100)
Holstebro		115 (21)	
Viborg		30 (5)	
General health, *n* (%)			
Excellent/very good	130 (32)	258 (47)	259 (44)
Good	149 (36)	209 (38)	243 (41)
Fair/poor	119 (29)	87 (16)	81 (14)
Missing	13 (3)		10 (2)
Well-being (WHO-5)			
Mean (SD)	61.3 (23.9)	70.6 (19.5)	68.5 (19.1)
Median (IQR)	64 (48–80)	76 (60–84)	72 (58–80)
Missing, *n* (%)	17 (4.1)	5 (0.9)	13 (2)
Social support for health (HLQ 4)			
Mean (SD)	3.3 (0.60)	N/A	3.3 (0.55)
Median (IQR)	3.4 (3–3.8)		3.4 (3–3.8)
Missing, *n* (%)	10 (2.4)		21 (3.5)
Ability to actively engage with healthcare providers (HLQ 6)			
Mean (SD)	3.6 (0.95)	N/A	3.8 (0.87)
Median (IQR)	3.8 (3–4.2)		4 (3.4–4.6)
Missing, *n* (%)	9 (2.2)		24 (4)
Understanding health information well enough to know what to do (HLQ 9)			
Mean (SD)	3.6 (0.97)	N/A	4.0 (0.82)
Median (IQR)	3.8 (3–4.3)		4 (3.6–4.6)
Missing, *n* (%)	9 (2.2)		24 (4)
Self-efficacy (GSES)			
Mean (SD)	27.4 (7.4)	N/A	29.3 (6.4)
Median (IQR)	29 (23–33)		30 (26–34)
Missing, *n* (%)	37 (9)		23 (4)
Patient activation[d], n (%)			
Disagree strongly/disagree	133 (32) [c]	N/A	41 (7)
Agree/agree strongly	264 (64)		533 (90)
Missing, *n* (%)	14 (3)		19 (3)
Patient activation[e], n (%)			
Disagree strongly/disagree	135 (33) [c]	N/A	85 (14)
Agree/agree strongly	264 (64)		486 (82)
Missing, *n* (%)	12 (3)		22 (4)

Abbreviations SD: standard deviation; IQR: interquartile range; HLQ: Health Literacy Questionnaire; GSES: General Self-Efficacy Scale; WHO-5: WHO-Five Well-Being Index.

[a] Based on the 411 (51%) patients who answered the research questionnaire; [b] Data from test 1; [c] Data from baseline; [d] I am confident that I can tell when I need to get outpatient care; [e] I am confident I can figure out solutions when new situations or problems arise with my health condition

STUDY I: DETERMINANTS OF REFERRAL TO PRO-BASED FOLLOW-UP

Study population

In all, 802 patients were included during the period from May 2016 to May 2018 (Figure 10) (120). The mean age was 49.3 SD (21.9) and 52% were men (Table 8). The original questionnaire-based analyses only included 411 responders (51%); however, all 802 patients were available for the register-based analyses. The patients' mean follow-up time in the study was 10.6 months (SD 6.6 months). By the end of follow-up (after 18 months or in January 2019), a total of 185 patients were referred to PRO-based follow-up, 172 patients had terminated outpatient care, and 52 patients had died (Table 9).

Figure 10 Flowchart of patients included in the study (120)

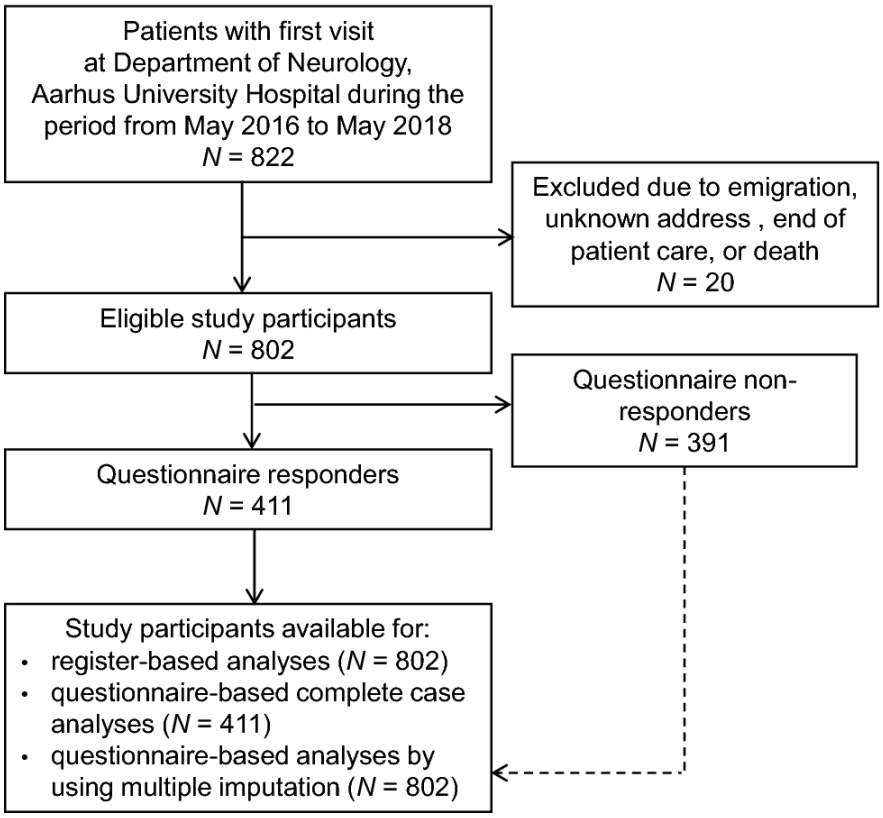


Table 9 Number of patients who had been referred to PRO-based follow-up, terminated outpatient care, or died at the three time points used in the study, *N*=802 *

	6-month follow-up	**12-month follow-up**	**18-month follow-up**
Referred to PRO-based follow-up	139	173	185
Terminated outpatient care	95	129	172
Died	26	43	52

* Fewer than 5 patients had emigrated at all three time points

Register-based determinants of referral to PRO-based follow-up

Compared to patients who lived with a partner/family, patients who lived alone had a lower probability of referral to PRO-based follow-up. Similarly, patients with a low level of education and household income had a lower probability of referral to PRO-based follow-up than patients with high levels. These results were consistent at all three measured time points. Among other variables associated with lower probability of referral to PRO-based follow-up were receiving temporary or permanent social benefits compared to self-supporting and having a psychiatric diagnosis compared to no diagnosis. These results were only statistically significant at 18-month follow-up. Moreover, men were more likely to be referred to PRO-based follow-up than women at 18-month follow-up. Patients with a medium level of co-morbidity were associated with lower probability of referral to PRO-based follow-up than patients with a low level of co-morbidity; however, this result was only statistically significant at 12-month follow-up. Referral to PRO-based follow-up was not related to age levels.

Questionnaire-based determinants of referral to PRO-based follow-up

Patients who reported a low level of perceived confidence to find solutions or resolve problems related to their health condition had a lower probability of referral to PRO-based follow-up than patients with high levels. This finding was consistent at all three time points. Furthermore, it was found that patients who reported a low level of perceived confidence to decide their need for outpatient care, well-being, health literacy (HLQ subscale 9: understanding health information well enough to know what to do), and general health had a lower probability of referral to PRO-based follow-up than patients who reported high levels or excellent/very good general health. These findings were consistent at two time points (at 12- and 18-month follow-up). Only at 18-month follow-up was a lower probability of referral found in patients with low self-efficacy compared to high self-efficacy. We also found that a 1-unit increase in mean scale scores of HLQ subscale 6 (ability to actively engage with healthcare providers) and HLQ subscale 9 increased the probability of referral to PRO-based follow-up at all three time points. In addition, 1-unit increase in mean scale scores of HLQ subscale 4 (social support for health) increased the probability of referral to PRO-based follow-up at 12- and 18-month follow-up.

Original data analyses and sensitivity analyses

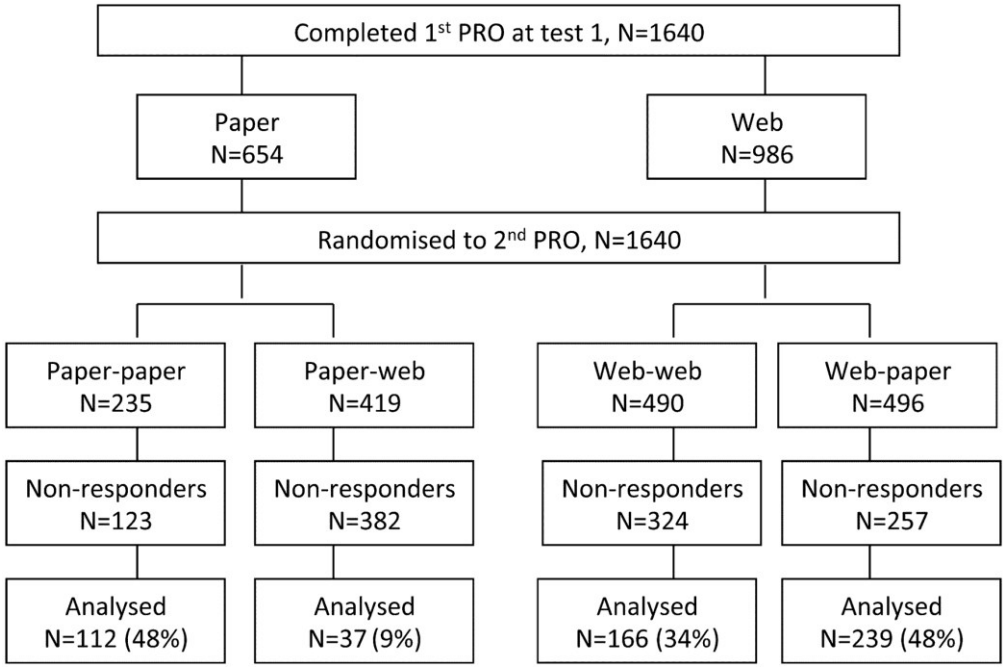
The original data analyses were consistent compared to the results based on multiple imputations. The results also remained consistent in the performed sensitivity analyses. Further details can be found in the supplemental materials of Paper I.

STUDY II AND STUDY III: TEST-RETEST RELIABILITY OF THE PRO-BASED ALGORITHM AND WHO-5

Study population

A total of 554/1640 (34%) patients completed the questionnaire at both test 1 and test 2. The mean age was 55.0 SD (16.6) years and 52% were men (Table 8). As shown in Figure 11, the response-rates varied across the four test-retest groups: 48% in the paper-paper and web-paper groups, 34% in the web-web group, and 9% in the paper-web group (142). Non-responders of the second questionnaire were younger, more likely paper-responders of the first questionnaire and reported a lower self-reported well-being and general health than responders. The median response time between test 1 and test 2 was 22 days, and the interquartile range (IQR) was 10 days. Missing values of the WHO-5 scale in either test 1 or test 2 were present for 14 patients, and these patients were not included in the analysis of Study III.

Figure 11 Flowchart of response method in test 1, randomization of response method in test 2, non-responders in test 2, and number of patients included in the analyses (142)



Agreement and test-retest reliability of the PRO-based algorithm

Perfect PRO-based algorithm agreement was observed in 82% (n = 454) and disagreement was observed in 18% (n = 100) of the test-retest cases (Table 10) (142). The estimated test-retest reliability of the PRO-based algorithm in terms of weighted kappa statistic was 0.67

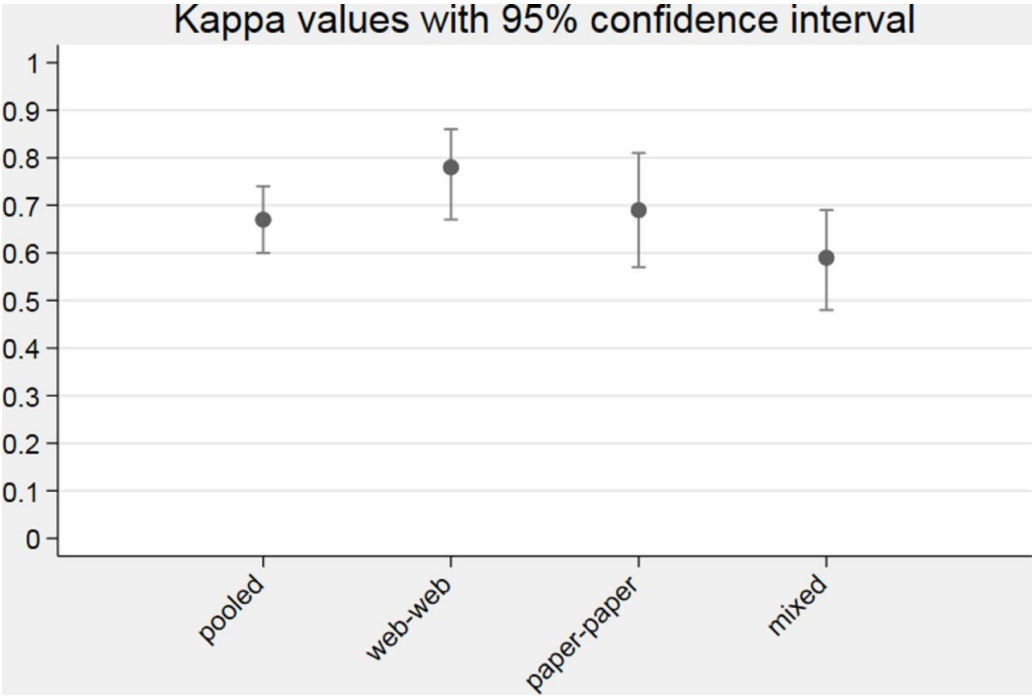
(95% CI 0.60 to 0.74). However, the kappa estimates varied across different methods of administration and were substantial in the web-web group; 0.78 (95% CI 0.67 to 0.86), but only 0.69 (95% CI 0.57 to 0.81) in the paper-paper group and 0.59 (95% CI 0.48 to 0.69) in the mixed-method group. Although the estimates varied between the groups, there were overlapping CIs (Figure 12) (142). The results did not alter markedly with restricted intervals between test 1 and test 2; however, a tendency towards a decrease of the kappa values was found if different methods of administration were used.

Table 10 Agreement between the automated PRO-based algorithms from test 1 to test 2 (142)

	PRO-BASED ALGORITHM TEST 2			
PRO-BASED ALGORITHM TEST 1	**GREEN (%)**	**YELLOW (%)**	**RED (%)**	**Total (%)**
GREEN	104 (19)	42 (8)	1 (0.1)	147 (27)
YELLOW	34 (6)	328 (59)	18 (3)	380 (69)
RED	0 (0)	5 (1)	22 (4)	27 (5)
Total	138 (25)	375 (68)	41 (7)	554 (100)

Green: No need of contact; Yellow: May need contact; Red: Need of contact

Figure 12 Test-retest reliability from test 1 to test 2 of the pooled PRO-based algorithms (n = 554), web-web (n = 166), paper-paper (n = 112), and the mixed groups (web-paper or paper-web, n = 276) (142)



Test-retest reliability of the items

The test-retest reliability parameters in terms of both unweighted and weighted kappa statistics of the single items in the epilepsy questionnaire (Appendix 2) were overall moderate to substantial. Moreover, in item categories within the framework of the PRO-based algorithm, kappa estimates were overall found to range from fair to substantial. In all items, missing values were under 5%; however, a ceiling effect occurred in the majority of the items, as we observed proportions of more than 15% at the upper ends of the scale.

Test-retest reliability and measurement error of the WHO-Five Well-Being Index

No systematic difference in the total WHO-5 scores between test 1 and test 2 was observed. The difference was 0.18 (95% CI −0.84 to 1.20). On the matter of measurement error, the Bland-Altman plot (Figure 13a) illustrates the 95% limits of agreement (143). The standard error of the measurement was 8.51 points (95% CI 8.03 to 9.05), which gave a MDC^{95} of 23.60 points (95% CI 22.27 to 25.10). The ICC was 0.81 (95% CI 0.78 to 0.84). The analysis was repeated across different methods of administration; however, this did not noticeably change the results (Figure 13b) (143). Nor did the results change with restricted intervals between test 1 and test 2. Cronbach's alpha estimates of the WHO-5 scale were nearly identical in test 1 and test 2: 0.89 (95% CI 0.87 to 0.90) and 0.89 (95% CI 0.87 to 0.91).

Figure 13a Bland-Altman plot of differences in WHO-5 score between test 1 and test 2 plotted against the mean, *N* = 540 (143)

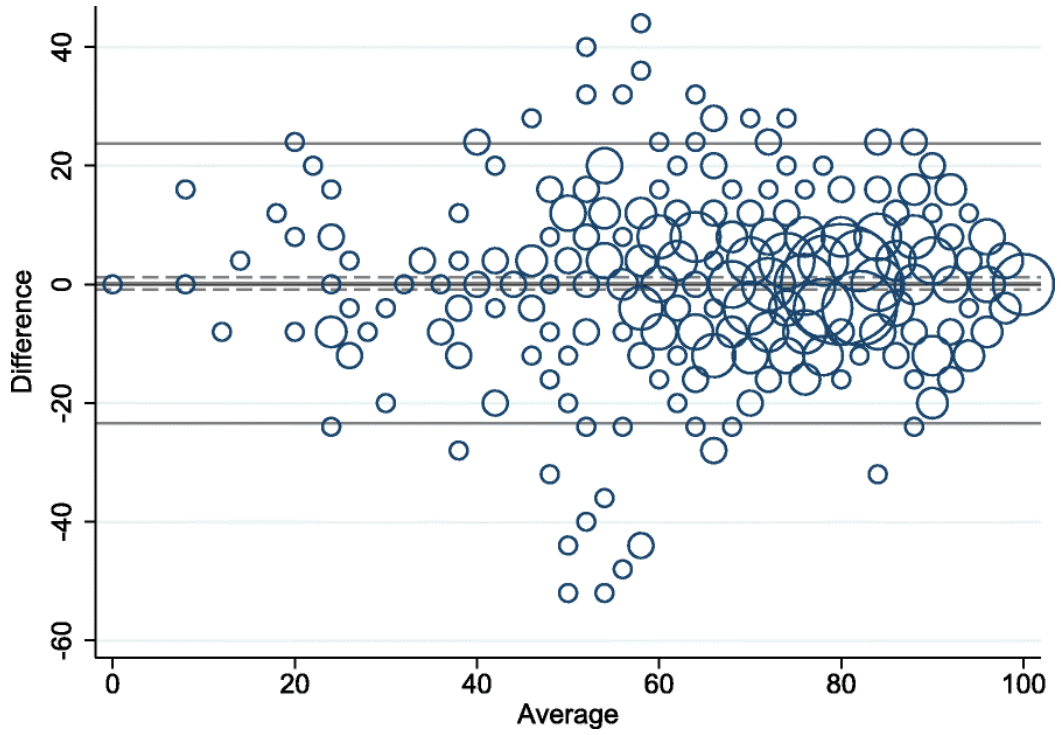
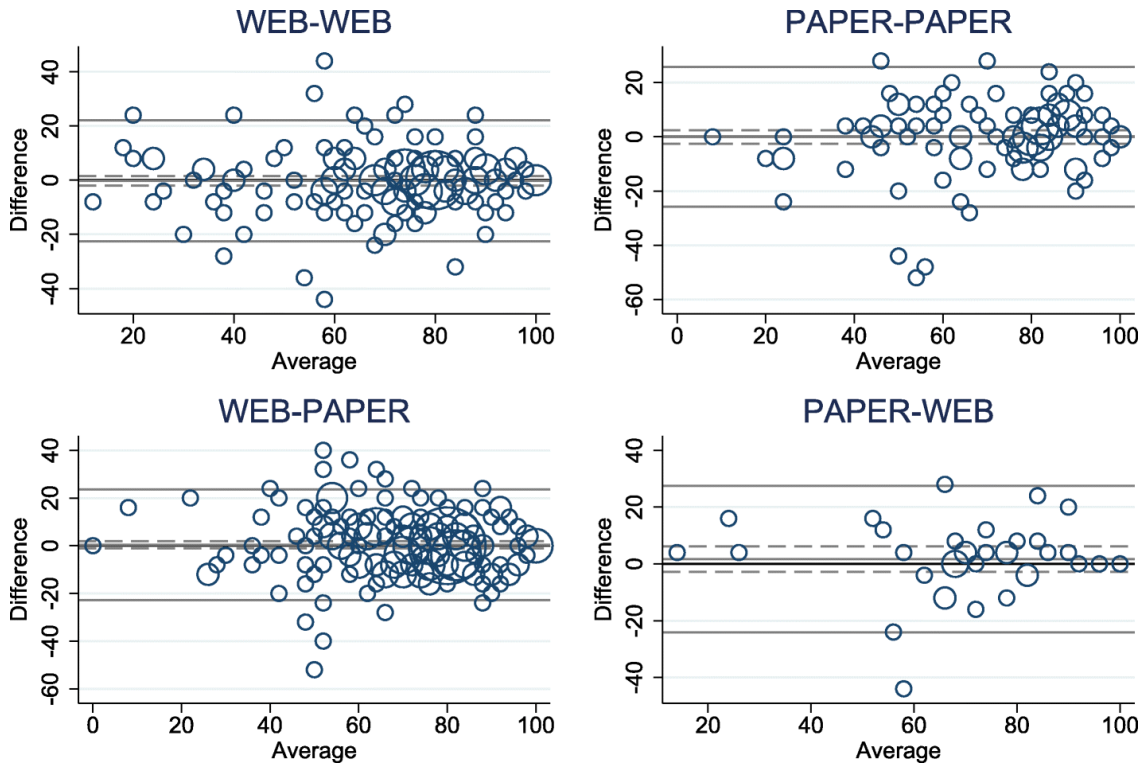


Fig. 13b Differences in the WHO-5 score between test 1 and test 2 plotted against the mean in the four methods of administration groups: web-web (*n* = 164), paper-paper (*n* = 107), web-paper (*n* = 233), and paper-web (*n* = 36) (143)

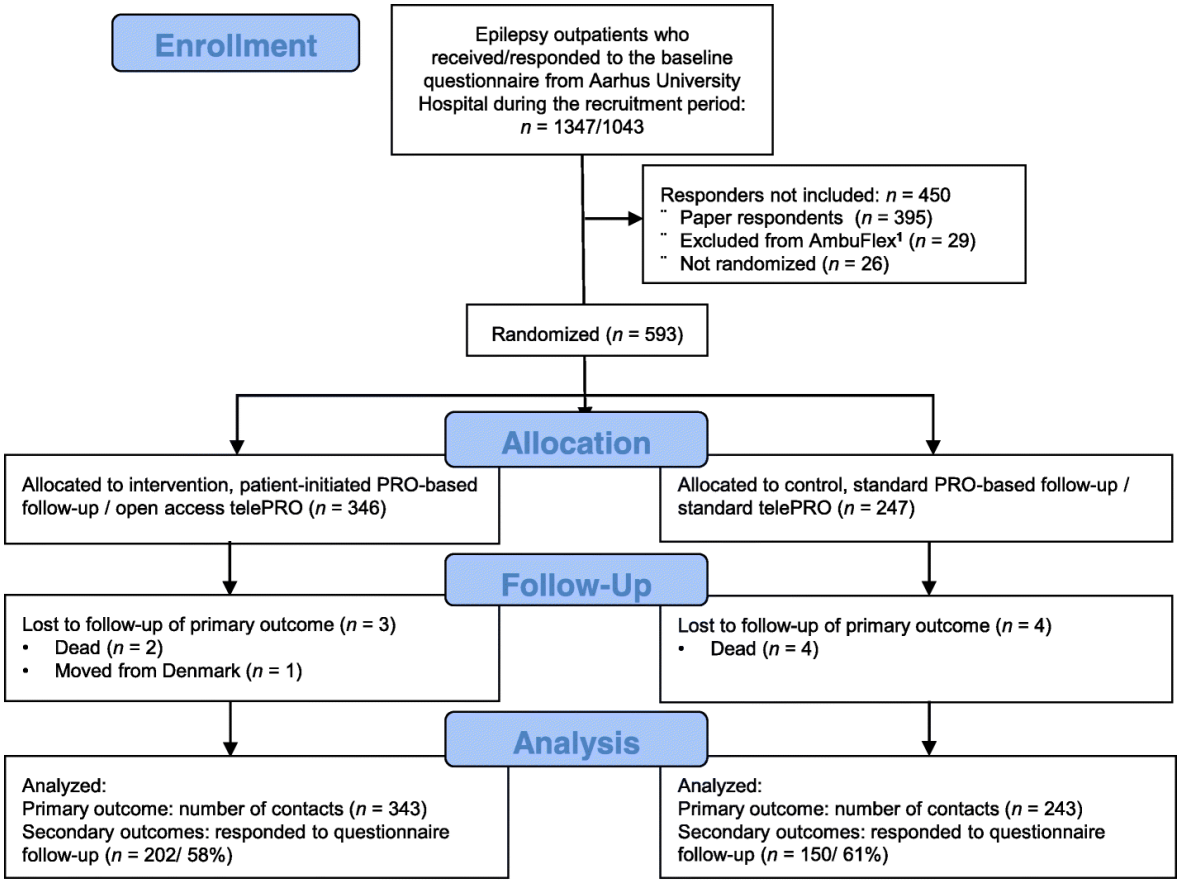


STUDY IV: THE EFFECTIVENESS OF PATIENT-INITIATED PRO-BASED FOLLOW-UP

Study population

A total of 593 patients were included in the study; 346 patients were randomized to receive the patient-initiated PRO-based intervention and 247 patients were randomized to the control arm (Figure 14) (144). The mean age of the 593 patients was 46.6 SD (17.2) years and 50% were men (Table 8). Patients in the two comparison arms were equally balanced with respect to baseline characteristics. Follow-up was 99% for primary outcome and 59% for self-reported secondary outcomes. A total of 43 patients were excluded in the per-protocol analyses because they declined to participate in the intervention arm.

Figure 14 The study CONSORT flow diagram (144)



[1] Patients who were excluded from standard telePRO after clinical assessment by a clinician of the patient's fixed-interval questionnaire response

Use of healthcare resources (primary outcome)

No statistically significant between-arm differences were found in the mean number of outpatient visits and telephone consultations (Table 11) (144). Patients in the intervention arm had a statistically significant, slightly lower number of emergency room visits, with an estimated difference of −0.11 (95% CI −0.21 to −0.01) compared to patients in the control arm. No statistically significant differences were found in the mean number of hospital admissions. Moreover, no statistically significant differences were found in the per-protocol analysis.

Table 11 Healthcare utilization during an 18-month follow-up period (144)

	Intention-to-treat population		
Primary outcome	**Intervention arm** ***n*** **= 343**	**Control arm** ***n*** **= 243**	**Mean difference (95% CI)**
Outpatient visits[a]			
Mean (SD)	0.45 (0.95)	0.42 (0.86)	0.03 (−0.11 to 0.18)
Median (Range)	0 (0–7)	0 (0–6)	
Telephone consultations[a]			
Mean (SD)	0.99 (1.88)	1.30 (2.46)	−0.32 (−0.68 to 0.05)
Median (Range)	0 (0–12)	1 (0–22)	
Hospitalizations[a]			
Mean (SD)	0.05 (0.29)	0.09 (0.49)	−0.04 (−0.10 to 0.03)
Median (Range)	0 (0–3)	0 (0–5)	
Emergency room visits[b]			
Mean (SD)	0.07 (0.38)	0.19 (0.72)	−0.11 (−0.21 to −0.01)
Median (Range)	0 (0–4)	0 (0–7)	

Abbreviations SD: Standard deviation; CI: Confidence interval

[a]at the Department of Neurology, Aarhus University Hospital; [b]at Aarhus University Hospital

Patient health outcomes and the patient perspective (secondary outcomes)

No statistically significant between-arm differences were found in number of seizures during last year, treatment side effects, and general health (Table 12) (144). Patients in the intervention arm had a statistically significant lower well-being; −3.21 (95% CI −6.38 to −0.05) than patients in the control arm. No statistically significant differences were found in health literacy, self-efficacy, patient activation, confidence, safety, or satisfaction. Moreover, no statistically significant differences were found in the per-protocol analysis.

Attrition and sensitivity analyses

The study population only included web responders; thus 395 paper responders were excluded (Figure 14) (144). Web responders did differ from paper responders: they were younger, better educated, and had a higher level of health literacy and self-efficacy. No differences were found in gender, general health, or well-being between web and paper responders. Approximately 60% of the participants in both arms responded to the research questionnaire at 18-month follow-up. Non-responders did differ from

responders, as non-responders were younger, less educated, and reported a lower level of well-being, general health, and health literacy (HLQ 6: Ability to actively engage with healthcare providers) at baseline. No differences were found in gender, cohabitation status, self-efficacy, disease activity and duration, or health literacy (HLQ 4: social support for health and HLQ 9: understand health information well enough to know what to do). The sensitivity analysis of the WHO-5 score indicated that the results could be biased in both directions based on four different scenarios. Further details can be found in the supplemental materials of Paper V.

Process outcomes

The number of logins to the intervention website 'My Epilepsy' and the number of completed questionnaires initiated by the included patients were generally low and decreased over time. Further information can be found in Paper V.

Table 12 Patient-reported outcomes measured 18 months after randomization (144)

	Intention-to-treat population		
Secondary outcomes	**Intervention arm** ***n*** **= 202**	**Control arm** ***n*** **= 150**	**Difference[a] at 18-mo. follow-up (95% CI)**
Well-being (WHO-5)			
Mean (SD)	66.99 (19.45)	69.29 (18.01)	−3.21 (−6.38 to −0.05)
Missing, *n* (%)	4 (2)	4 (3)	
Social support for health (HLQ 4)			
Mean (SD)	3.24 (0.60)	3.38 (0.53)	−0.08 (−0.17 to 0.02)
Missing, *n* (%)	5 (2)	4 (3)	
Ability to actively engage with healthcare providers (HLQ 6)			
Mean (SD)	3.84 (0.82)	3.87 (0.89)	−0.05 (−0.21 to 0.10)
Missing, *n* (%)	6 (3)	4 (3)	
Understanding health information well enough to know what to do (HLQ 9)			
Mean (SD)	4.03 (0.77)	3.97 (0.85)	0.009 (−0.13 to 0.15)
Missing, *n* (%)	6 (3)	4 (3)	
Self-efficacy (GSES)			
Mean (SD)	29.78 (5.69)	29.73 (6.14)	−0.22 (−1.22 to 0.78)
Missing, *n* (%)	7 (3)	4 (3)	
General health			
Mean (SD)	2.63 (0.93)	2.60 (0.82)	0.05 (−0.10 to 0.19)
Missing, *n* (%)	1 (0.05)	1 (0.07)	
No. of seizures last year			
Mean (SD)	2.50 (11.89)	3.20 (10.21)	−0.72 (−3.20 to 1.75)
Missing, *n* (%)	36 (18)	28 (19)	
Side effects			
Mean (SD)	1.54 (0.76)	1.56 (0.83)	−0.03 (−0.18 to 0.11)
Missing, *n* (%)	6 (3)	1 (0.07)	
Patient activation[b]			
Mean (SD)	3.42 (0.65)	3.34 (0.77)	0.04 (−0.10 to 0.17)
Missing, *n* (%)	6 (3)	4 (3)	
Patient activation[c]			
Mean (SD)	3.22 (0.72)	3.12 (0.75)	0.01 (−0.13 to 0.16)
Missing, *n* (%)	5 (2)	4 (3)	
Confidence			
Mean (SD)	1.39 (0.65)	1.33 (0.53)	0.03 (−0.9 to 0.16)
Missing, *n* (%)	21 (10)	9 (6)	
Safety			
Mean (SD)	1.41 (0.70)	1.35 (0.56)	0.02 (−0.12 to 0.16)
Missing, *n* (%)	37 (18)	14 (9)	
Satisfaction			
Mean (SD)	1.63 (0.68)	1.61 (0.59)	0.01 (−0.13 to 0.15)
Missing, *n* (%)	35 (17)	18 (12)	

Abbreviations SD: Standard deviation; CI: Confidence interval; WHO-5: WHO-Five Well-Being Index; HLQ: Health Literacy Questionnaire; GSES: General Self-Efficacy Scale

[a]Adjusted for the baseline measure; [b]I am confident that I can tell when I need to get outpatient care; [c]I am confident I can figure out solutions when new situations or problems arise with my health condition

DISCUSSION

In this section the key results for each study are discussed. In addition, a critical evaluation of methodological aspects are discussed separately for Study I and Study IV, and combined for Studies II & III.

DISCUSSION OF RESULTS

Study I: Determinants of referral to PRO-based follow-up

The results from our study indicate that PRO-based follow-up is offered to a selected group of socioeconomically advantaged patients with epilepsy. PRO-based follow-up is a new model of care and to the best of my knowledge, no other studies have investigated associations between sociodemographic, personal, or disease-related determinants and referral to PRO-based follow-up. Therefore, this study is explorative where the results can provide hypothesis-generating evidence indicative of a potential association between a specific factor and the outcome (145). Since the essence of PRO-based follow-up is that patients regularly fill in questionnaires, I found it relevant to consider studies of factors associated with questionnaire non-response. Results from such studies show that lower socioeconomic status, male gender, younger age, solo living, and poorer health status are associated with questionnaire non-responses (92-98,146). In our study, we found an association between a lower level of education, household income, and solo living and non-referral to PRO-based follow-up. However, we did not find an association between younger age and referral to PRO-based follow-up. This may be related to study differences, as our study did not investigate factors associated with questionnaire non-response in PRO-based follow-up but to referral to PRO-based follow-up decided in the clinical encounter. In line with our findings, other studies have also found an association between lower probability of participation in self-management interventions if patients had a lower level of education (147) or had lower self-efficacy (148). Furthermore, a study found that referral to advanced heart therapy was associated with higher patient activation among patients selected for therapy compared to patients who were not selected (149). A Danish study found that well-educated patients with endometrial cancer more often initiated medical attention if symptoms of recurrence occurred (150). Thus, there seems to be a selection of patients with regard to participation and use of different initiatives in the healthcare system.

The healthcare system seeks for solutions that aim to manage the balance between resources used for acute care management and the needs of patients with chronic diseases (10,151). Self-management interventions may be a means to bridge the gap between patients' needs and healthcare capacity, as they have been shown to give patients knowledge and skills to manage chronic diseases (147) and have the potential to reduce the use of healthcare resources (44). However, successful patient self-management requires sufficient ability to manage the consequences inherent in living with a chronic disease (151). The patients' capacity should balance the burden of disease and treatment, as low capacity and high disease and treatment workload may inhibit self-management, treatment adherence, and health outcomes (152). This may support the need for

differential follow-up activities in the care for patients with chronic diseases. PRO-based follow-up is one solution that aims to offer individual and flexible remote follow-up and optimize the use of healthcare resources, avoiding unnecessary in-clinic visits if patients are well-treated (4). The results from our study indicate that PRO-based follow-up is mainly offered to a selected group of socioeconomically advantaged patients with epilepsy, while the less advantaged patients remain with the traditional follow-up method. In a qualitative study, clinicians experienced that patients seen in-clinic have more complex problems after implementation of PRO-based follow-up (153), which clearly demonstrates that use of PRO-based follow-up means that resources can be reallocated for use in patients with a high symptom burden.

A Danish qualitative study in patients with rheumatoid arthritis found that some patients valued the autonomy and independence of remote PRO-based follow-up, while others preferred face-to-face contact (154), which was also the case in PRO-based follow-up among outpatients with epilepsy (155). These findings support the notion that this type of follow-up is not optimal for the whole outpatient population, but referral to PRO-based follow-up must be based on an individual clinical decision and the patient's preferences (4,155). However, since no standardized guidelines were defined regarding which patients the clinicians should refer to PRO-based follow-up, the personal preferences among the clinicians probably played a role when deciding to refer or not. Two qualitative studies regarding the use of PRO-based follow-up in outpatients with epilepsy have identified barriers related to participation among both patients and clinicians (153,155). For clinicians, the lack of interpersonal contact was a negative consequence of remote PRO-based follow-up (153). Some clinicians felt unsure about some of the patients' capabilities to participate in PRO-based follow-up even though the patients had already been referred. Another qualitative study found that clinicians perceived health literacy to be a potential barrier for patients completing PRO measures (156). For patients, participation in PRO-based follow-up could result in a sense of rejection and disconnection from the clinic, and some patients felt a lack of confidence in their own ability to participate in PRO-based follow-up (155). These findings underline the need for shared decision-making between the patient and the clinician before referral to PRO-based follow-up (153,155).

The use of PRO data in clinical practice has the potential to enhance patients' capacity for self-management (157) by, for example, improving communication (39,41,43,44,46-50), increasing awareness of psychosocial problems (43,44,46,47,49), and increasing the patients' understanding of the disease (36,44,154,155,158). In our study, we found that only a selected group of socioeconomically advantaged patients are able to obtain these benefits when PRO data are used in remote follow-up. Therefore, to prevent health inequality, it is relevant to pay attention to how PRO data can be used in outpatient

follow-up among vulnerable patients. For example, PRO data could be used prior to telephone consultations or in-clinic appointments. Moreover, proxy solutions could be used among patients with cognitive disabilities who are unable to complete a questionnaire by themselves.

Study II: Test-retest reliability of the PRO-based algorithm

Our study is the first reliability study of an epilepsy questionnaire coupled with a PRO-based algorithm used to identify patients who need clinical attention. The questionnaire and PRO-based algorithm are now used among outpatients with epilepsy at hospitals in all five Danish Regions. Overall, we found that test-retest reliability in terms of kappa values and percentage agreement was moderate to substantial for both the PRO-based algorithm and the items included in the questionnaire. However, kappa values are influenced by skewed distribution, number of classes, and systematic differences between the two measurements (22,159). Thus, it is recommended to present study results in cross tables in addition to giving the kappa values. A cross table such as Table 10 was only used to present results regarding agreement of the PRO-based algorithm. A skewed distribution leads to less room for real agreement due to a higher fraction of chance agreement (22). In our data, we saw a skewed distribution in several items in the epilepsy questionnaire because it was filled in by a homogeneous population with stable disease and low symptom burden. This was also supported by the finding that the PRO-based color algorithm gave only a few red responses. This skewed distribution may have resulted in an underestimation of the kappa values in our study.

We found the web- and paper-based methods of administration to be equivalent, and this is supported by other studies (28,30-32). According to the ISPOR's guideline, it is important to assess the comparability of web- and paper-based versions of a questionnaire (29). In 2015, mixed-mode administration methods in terms of use of both web- and paper-based questionnaires were nearly equally balanced in the AmbuFlex system; however, in recent years, the web-based questionnaire has taken over and in 2019, only 4.3% of the questionnaire responses were paper-based (75). We expect this process to continue, as it follows the dissemination of web-based solutions in Danish society (160). In light of this, our study aim regarding whether the mixed administration methods would influence our results may not have the same relevance in the future. We found, contrary to other studies (30,161-164), a tendency toward lower equivalence between the two methods if they were used interchangeably. This may be related to our study design in which patients could select their preferred mode of administration at the first test and were thereafter randomized to a compulsory mode of administration at the second test. This probably caused both a low response rate of only 9% in the paper-web group and perhaps also affected the patients' habits and reflection in items of the second test, which may have caused a lower equivalence between the two mixed methods.

Among the three main measurement property domains in the COSMIN framework, we only investigated reliability. A measurement can be reliable without being valid, but cannot be considered valid if it is not reliable (22). Content validity is considered one of the most important measurement properties of a PRO measure (165). During the development process of the epilepsy questionnaire, we attached importance to the content validity aspect, since the relevance of a PRO measure to both patients and clinicians is pivotal to their utility in clinical practice (4). The validity of some of the constructs in the questionnaire has been documented, such as the WHO-5 scale (87,88), items from SF-36 (89,90), and items from the SCL-92 (91). However, some of the items were self-constructed, and the validity of these items has not been documented. Moreover, the validity of the PRO-based algorithm has also not been documented, and this should be an area of investigation in future research.

Study III: Test-retest reliability and measurement error of WHO-5

Our study is one of the first studies to investigate the reliability of the WHO-5 scale, not only including internal consistency but also test-retest reliability and measurement error. Internal consistency of the WHO-5 scale has been well documented in several studies across many patient populations and countries. We found a Cronbach's alpha of 0.89, which is in line with other studies that have reported estimates that ranged from 0.82 to 0.95 (99-103,166-170). However, only two studies regarding test-retest reliability of the WHO-5 scale have been identified (171,172). A study by Bonnin et al. evaluated the test-retest reliability of the Spanish version of the WHO-5 scale in a euthymic patient population with bipolar disorder (171). They compared the WHO-5 score at baseline and after 10 days of the first administration. The reliability of the WHO-5 scale was reported to be high, with a Pearson correlation coefficient of 0.83, and no statistically significant changes were found. However, they did not report measurement error, and the result was based on a small sample size of only 16 patients (171). Another study from Germany by Englbrecht et al. has also assessed the test-retest reliability of the WHO-5 scale among patients with rheumatoid arthritis (172). The correlational estimate of the retest reliability was 0.67, which was below the a priori criterion of a correlation of $\rho \geq 0.70$, and again a measurement error was not reported. The WHO-5 scale was measured at two time points with a range of 10–14 weeks between the two measurements. Only patients without depression and patients who had not received antidepressant treatment between the two measured time points were included in the analysis, but the actual number of patients was not reported (172). The findings from these two studies are not directly comparable to our study, as we used ICC to measure the test-retest reliability estimates in a different patient population, thus there is a need for further research.

Issues related to reliability of PRO measures at the individual level compared to group level have been raised in the literature (22,173). A reliability coefficient of 0.70 is considered adequate for use of measures at group level, whereas for individual level use a minimum threshold of 0.90 has been suggested (22,173). However, it depends on how the measurement instruments are used in a population. For diagnostic and prognostic purposes, the intention is to distinguish between different levels or courses of diseases; hence, reliability parameters are relevant. However, if the intention is to evaluate changes in health status, parameters of measurement error are relevant (22). In PRO-based follow-up, the WHO-5 is used to identify risk of depression among outpatients with epilepsy. In our study, the ICC of the WHO-5 scale ranged from 0.78 to 0.84; hence it was not above the recommended threshold of 0.90. The ICC estimates in our study were probably to some extent underestimated due to a homogeneous study population. However, clinicians should be aware of the scale's limitations and how to interpret the scale when using WHO-5 at the individual patient level.

To the best of my knowledge, our study is the first to evaluate the measurement error of the WHO-5 scale. The large measurement error in terms of minimal detectable change (MDC) found in our study is an interesting finding if the WHO-5 scale is used to measure change in well-being over time, for example, to measure the effect of an intervention in a RCT. However, a change can be statistically significant though not clinically important or relevant to the patients (22). The term minimal important change (MIC) is related to the smallest change in the score that is considered to be important by the patients. Therefore, the MIC can be used to look for clinically relevant changes in a RCT (22). To the best of my knowledge, an MIC has not previously been estimated for the WHO-5 scale; however, an important change has been reported to be at least 10 points (88). We observed an MDC of 23 points. If the MDC is larger than the MIC, it is not possible to determine whether the change is an important, clinically relevant change or whether it is due to measurement error (22). As a result of this potential limitation of the WHO-5 scale, the use of WHO-5 as an outcome to measure change over time should be carefully considered.

Study IV: The effectiveness of patient-initiated PRO-based follow-up

We did not identify a difference in healthcare use between patient-initiated PRO-based follow-up and fixed-interval PRO-based follow-up. This is not in line with results from RCT studies that have shown that patient-initiated follow-up has the potential to reduce healthcare use compared to that used in control groups with conventional follow-up in patients with rheumatoid arthritis (174), inflammatory bowel disease (140,175), and endometrial cancer (70). Furthermore, in patients with epilepsy, a retrospective evaluation showed that an open access model of healthcare delivery using telephone consultations could prevent in-clinic appointments (176). As hypothesized, we did not find any differences in clinical outcome measures between the intervention arm and the control

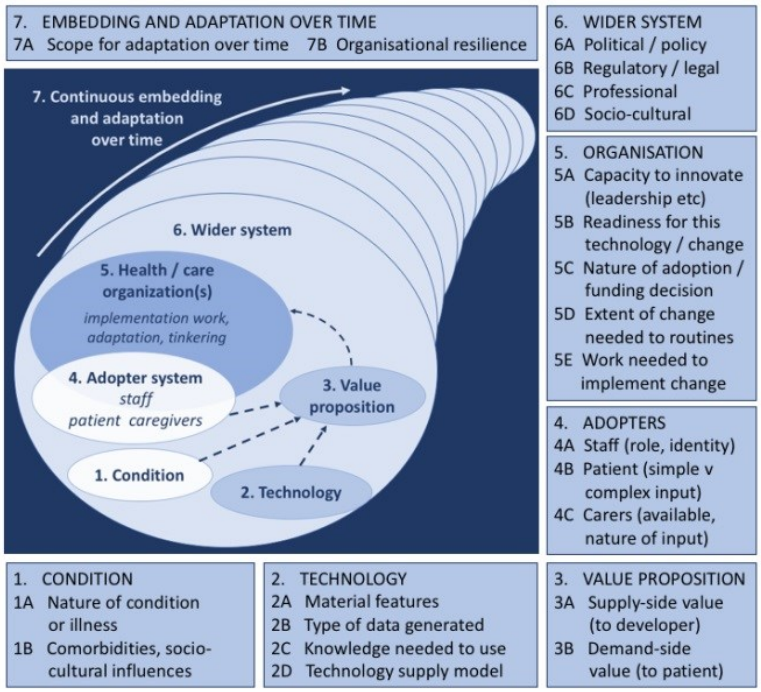
arm in our study. This finding is in line with results from RCT studies evaluating the effect of patient-initiated follow-up in patients with rheumatoid arthritis (174), inflammatory bowel disease (175), and chronic obstructive pulmonary disease (177). Furthermore, a systematic review of RCT studies evaluating patient-initiated interventions found no differences in HRQOL and psychological outcome measures compared to traditional follow-up (73). In the intervention arm, we found a statistically significant lower self-reported mental well-being of 3 points. However, since a clinically relevant change on the WHO-5 scale is considered to be at least 10 points (88), the difference was probably not clinically significant.

Contrary to our study, the control groups in the above-mentioned RCT studies were offered fixed-interval in-clinic visits at the department and not fixed-interval questionnaires as the basis for follow-up. Therefore, the difference in number of contacts between the two arms in our study was probably smaller than in the other studies. When designing the study, we discussed the opportunity to include a control arm with scheduled in-clinic visits. However, as our study was a pragmatic RCT, we had to customize the intervention to real-world clinical practice, and a control group with in-clinic visits was not possible for the following reason. The Department of Neurology at Aarhus University had already implemented AmbuFlex with fixed-interval PRO-based follow-up in early 2012, and in 2016, this was considered standard care for approximately 2500 outpatients with epilepsy. Hence, the target group of patients to be included in our study was primarily offered fixed-interval PRO-based follow-up as standard care and not scheduled in-clinic visits. To test the effectiveness of patient-initiated PRO-based follow-up versus in-clinic visits follow-up, we could have decided to include incident patients. However, a single unit study may have been less feasible, as it would have required a very long inclusion period.

Non-adherence to health technologies by their intended users is a common problem (178). We collected process outcomes during follow-up, and our data indicated that the patients did not use the intervention website very often, and usage decreased during follow-up. It is possible that some patients found it easier to call the clinic instead of using the web-based intervention. On the other hand, it could also be explained by a lack of a need for clinical attention. Greenhalgh et al. developed a theoretical framework regarding non-adoption, abandonment, scale-up, spread, and sustainability (NASSS) of healthcare technologies (Figure 15) (179). The framework can help address key challenges in seven different domains: 1) the condition, 2) the technology, 3) the value proposition, 4) the adopters, 5) the organization, 6) the wider context, and 7) interaction and adaptation between all these domains over time (179). The framework can be related to both the design and development phases of a new technology or in a scale-up of technology programs, but can also be used retrospectively to explain program failures (179). In the

following, potential key challenges related to usage of our intervention will be discussed based on this framework.

Figure 15 The non-adoption, abandonment, scale-up, spread, and sustainability framework from Greenhalgh et al. (179)



The condition: The included patients were already attending fixed-interval PRO-based follow-up. In this follow-up method, many patients have a stable disease activity, and need of clinical attention might not be necessary for a prolonged period. Before enrollment in the study, a clinician had evaluated the patient's capabilities regarding participation in fixed-interval PRO-based follow-up, and the patient was involved in this decision.

The technology: The AmbuFlex system was implemented at the department in 2012; hence, the technology system was well-integrated and used in everyday practice before the start of the study. However, due to data security, patients needed to login to a national eHealth Portal (https:\\Sundhed.dk) to access the intervention website 'My Epilepsy'. This process involved some extra login steps compared to accessing the questionnaire in the control arm. In the case of technical problems, external support from Sundhed.dk was needed. Technical issues occurred a few times during follow-up, but the problems were solved as quickly as possible and often within a couple of days.

The value proposition: The intervention aimed to reduce the use of healthcare resources and provide patient-centered care, which was desirable for both patients and the healthcare system. Results from RCT studies indicate that patient-initiated interventions are safe and

do reduce healthcare use (5,6,73). Both patients and clinicians were involved during the intervention development to increase the utility. During this process, some patients saw the value of selecting the time and content of their own individual follow-up as well as having access to previous questionnaire responses.

The adopters: For patients, the intervention demanded some skills related to self-management, as the patients had to decide their own need of contact with the healthcare system. Some patients may appreciate this autonomy, while others may value security and prefer health professionals deciding for them (154,155). Furthermore, as pointed out by May et al., a shift in healthcare from clinic to home places new demands on the patient, hence, a growing burden for some patients (180). A higher degree of responsibility regarding self-management could be perceived as a burdensome task for some patients who did not have the capacity to meet these demands, e.g. self-referring when they needed to be seen by a clinician. Patients received only written information about the intervention, and we assumed that the patients agreed to and understood the intervention if they did not decline to participate. The introduction may have been improved by supplemental initiatives such as a brief informative film about the intervention and an interactive dialogue with a clinician. This would be more time-consuming, but the included patients may have felt better informed about their active role during follow-up. For clinicians, the intervention did not demand that the staff learned new skills. They handled the incoming intervention PRO responses in the same web system as was used in the control arm.

The organization: The organizational readiness seemed to be sufficient; as the department had a highly effective leadership who was engaged in the development and implementation of a new patient-centered intervention. However, the work related to implementation of the intervention may have been underestimated, as the technology used was already implemented at the department. A shared vision of the intervention's potential compared to fixed-interval PRO-based follow-up may have been unclear to some in the local clinical team.

The wider context and adaptation: A system-wide shift to patient-initiated PRO-based follow-up and ongoing adaptation in an outpatient epilepsy population has not taken place after the project ended. The department continues using fixed-interval PRO-based follow-up, and the patients can always call the department between the scheduled questionnaires if needed.

In summary, by using the NASSS framework retrospectively, challenges were identified primarily in the adopter domain. Non-adherence in our study may have been related to the complexity of the patients' tasks in patient-initiated follow-up, which demanded that the patients had to both judge their need of and initiate contact. Patient-initiated PRO-based follow-up should be considered as a complex intervention in the healthcare system.

Complex interventions have to be evaluated to identify whether they are effective in everyday practice (181). The Medical Research Council's framework regarding development and evaluation of complex interventions consists of four iterative phases: development, feasibility and piloting, evaluation, and implementation (181). It may have been helpful if we had used this framework or the NASSS framework prospectively in our study, thereby enabling us to address some of the key challenges.

CRITICAL EVALUATION OF THE METHODS

The methodological quality of a study is important, as the quality indicates the trustworthiness of the findings. In research, bias can occur if any systematic flaws in the design or conduct of a study distort the findings (182). Bias can either underestimate or overestimate the association or the intervention effect. In this section, evaluation of internal validity and risk of bias is supported by standardized guidelines according to the study designs (183-185). In Study IV, other methodological aspects are also discussed, such as ethical considerations and selection of self-reported outcomes. Finally, generalizability (external validity) of the results is discussed.

Study I: A prospective cohort study

Observational studies are well-accepted in the study of risk factors and prognosis (186). However, as in all observational studies, systematic errors and lack of methodological quality could affect the results in our study. The Quality In Prognosis Studies (QUIPS) tool was used in the critical evaluation of the applied method and to assess risk of bias (183). The following aspects are discussed below: selection bias (study participation and attrition), information bias (determinants and outcome measurement), confounding, and statistical analysis and reporting.

Selection bias

The enrollment of patients in the study is considered to be without bias, because patients were consecutively enrolled during a 2-year period. Eligible patients were identified in a regional register based on four selected ICD-10 codes. The code selection was based on a review of ICD-10 codes in a random sample of outpatients with epilepsy attending PRO-based follow-up at Aarhus University Hospital from January 2015 to March 2015 and given advice from a clinical epilepsy expert. Data from the random sample showed that the four selected ICD-10 codes covered 96.4% of the diagnoses.

The study had complete register-based data regarding several determinant variables and the outcome in all included participants. Hence, selection bias in the register-based results is considered to be limited. However, this was not the case in questionnaire-based data, as the response rate of the research questionnaire was only 51%. The low response rate could

lead to biased results in both directions. Non-responders did differ from responders, and non-responders were less likely to be referred to PRO-based follow-up and, hence, would be related to the outcome. Multiple imputation (MI) combined with sensitivity analysis is considered to be one of the most reliable statistical methods for handling missing questionnaire data (187). Therefore, we used MI to deal with the missing data problem based on the assumption that data were missing at random (MAR). We took advantage of using register-based data in the MI model, since we had access to register-based data of all patients including the non-responders. We also performed sensitivity analyses by changing the variables used in the MI model and by changing the imputed values, as the MAR assumptions could not be validated. These analyses did not change the associations markedly, and therefore support unbiased estimates.

Information bias

The quality of register-based data highly depends on the completeness and validity of the registers used (188). The data quality in the regional Hospital Business Intelligence (BI) Register has not been documented; however, Central Denmark Region points to the importance of the valid data that are used for several purposes in the Region (106). We obtained data about ICD-10 codes, emigration, and death, and these data are considered to be of high quality in the BI Register. The validity of the ICD-10 codes has not been documented in the BI Register, but the register monthly submits standardized data to the Danish National Patient Register (DNPR). The validity in terms of positive predictive value for the epilepsy diagnoses in the DNPR has been estimated to be 81.4% (95% CI 75.2 to 86.3) (189). Lack of registration of ICD-10 codes could potentially occur, but the registration is based on administrative requirements; hence, the lack is considered to be random and will not lead to bias. We also obtained data about patients who had stopped attending outpatient care, and the quality of this data depends on the procedure the department uses to register this type of data. We only had access to a registered date and it could thus be possible that the number of patients who ended outpatient care during follow-up was underreported; however, I do not believe this has influenced our findings.

According the national registers, information in the Danish Civil Registration System is used continuously for administrative purposes, and high quality is ensured through ongoing validation of the data recorded (114). Further, the education registers are of high validity; however, Statistics Denmark reports misclassification ranging from 0–3% up to 10% (116). The income data from the Danish Registers in Income and Transfer Payments are assumed to equal the real income and are of high quality. However, data of direct importance to the administrative authority are assumed to be more reliable than data without that kind of importance (117). Data in the Danish Register for Evaluation and Marginalization (DREAM) is generally of acceptable quality. Agreement between data in DREAM and self-reported information has been found to be highest for self-supporting

and retirement pension and poorest for labor market-related benefits (118). Misclassification in the DNPR is primarily related to changes of codes over time and multiple definitions used in the system. Data are assumed complete since 2000, as since then the DNPR formed the basis for payment to public hospitals (119). We used historical data related to co-morbidity before patients were included in the study, and misclassification of the ICD-10 codes is considered to be limited. Overall, any misclassification of register-based determinants is considered to be non-differential, which may have biased our results toward the null. This assumption is based on the premise that the misclassification was random, as the data collection in the registers is based on administrative requirements. Moreover, any misclassification took place before referral to PRO-based follow-up and, hence, would not be related to the outcome.

Misclassification of the outcome, referral to PRO-based follow-up, is considered to be limited. This information was also register-based and retrieved from the AmbuFlex database (4). The system is used in daily clinical practice, and referral is registered by a clinician in the system. Potentially, the clinician could forget to register the patient in the system; however, I believe this registration error to be of minimal importance to our study. The clinicians were not involved in the data collection in the study; thus, any misclassification was not related to the determinants.

Self-reported bias could be related to, for example, social desirability or recall bias (190). Some of the constructs of interest could be private and sensitive topics for some patients. For example, patients with low health literacy may have overestimated their exact level. In all questions, patients respond to their current status, except for the WHO-5 scale, in which the patients were instructed to consider mental well-being within the previous 2 weeks. This could have been confusing for some patients. However, the participants filled in the questionnaire before referral to PRO-based follow-up took place. Therefore, any misclassification of self-reported data probably resulted in non-differential bias, which may have biased our results toward the null.

Confounding

Potential confounding variables were selected a priori based on previous studies regarding factors associated with questionnaire non-response (92-98). Thus, we assumed that these factors also could be associated with PRO-based follow-up. The confounder variables were register-based and nearly complete with few missing values. In addition, the confounder variables were included in a statistical regression model. However, we do not know whether inclusion of other potential confounding variables would have influenced our results. For example, we did include co-morbidity, but it may have been relevant to include epilepsy disease severity as a confounder variable as well; however, we did not have access to the patients' medical records to gather this information. If

disease severity was a confounder in the associations in our study, these associations may have been overestimated due to lack of adjustment.

Statistical analysis and reporting

The choice of statistical model is considered adequate, and data have been presented in a transparent way. All analyses were based on a pre-defined statistical analysis plan; thus, no selective reporting of the results took place. Furthermore, we have used robust statistical techniques (multiple imputation) and sensitivity analyses to investigate the potential impact of bias, especially related to missing self-reported data.

Overall, bias of our findings is considered low in all aspects except regarding missing self-reported data. I do not, however, expect this selection bias to have had an impact on the internal validity of our study.

Study II and Study III: Test-retest reliability studies

A test-retest design with at least two measurement time points is recommended in the evaluation of reliability and measurement error properties (22). However, lack of methodological quality could have affected the results in our study. We followed the recommended COSMIN guidelines regarding risk of bias related to reliability and measurement error (184,191). Two aspects that could bias our findings, selection bias and design requirements, are discussed below.

Selection bias

The enrollment of patients in our study is considered adequate, as patients attending PRO-based follow-up from three departments were consecutively enrolled during a 9-month period. However, the response rate at the second measurement time point was only 34%; thus, selection bias cannot be ruled out. The differences between responders and non-responders of the second test indicated that the responders were a healthier group of patients than the non-responders. This could potentially underestimate the reliability parameters due to the homogeneous study population. The measurement error, however, was affected to a less extent since standard error of the measurement is a parameter of the measurement error reflecting the within-patient variation (134).

Design requirements

Two of the design requirements stated in the COSMIN checklists are to ensure that the patients are stable with respect to the construct under investigation in the interim period, by using an appropriate time interval between the two measurements (130,191). Based on clinical experiences, we assumed that outpatients with epilepsy had stable disease activity between the two time points. However, some data indicated a worsening of health status in the patients' second response, and we allowed for quite a long time interval for some patients: the maximum range was 104 days and the median range was 22 days. If the

study population was not stable due to fluctuating disease activity, this could have reduced the reliability parameters and made the PRO measures look unreliable even though it may accurately detect responsiveness (192). Assessing no real change in the measured constructs could be based on measurement of similar well-known reliable constructs at the same time or a question regarding change in health status included in the second questionnaire; however, the latter may induce the risk of recall bias. We performed sensitivity analysis, including only patients with a shorter interval, and we found a tendency toward increased reliability if the same method of administration was used and decreased reliability if two different methods of administration were used. A third requirement of the COSMIN checklist is to ensure similar test conditions (130,191). However, our aim was to investigate whether the methods of administration influenced the results. The patients completed the first questionnaire according to their preferred method of administration, but for some patients a compulsory change in administration method took place with regard to filling in the second questionnaire. Therefore, it is possible that the decrease in reliability found in the sensitivity analyses of the mixed methods was caused by this test-retest variation, rather than by unstable disease activity among the participants.

Other requirements stated in the COSMIN checklist, such as sample size and statistical methods, are considered to be sufficient in this study (130,191).

Overall, bias is considered to have played a moderate role in our findings due to the homogeneous study population, lack of clarity regarding change in health status between assessments, and environmental disruption. These aspects may have underestimated the reliability estimates in our study.

Study IV: A pragmatic randomized controlled study

The RCT design is considered to be the standard method for evaluation of effectiveness of a therapy or other interventions to improve health outcomes (186). We conducted a pragmatic RCT which involves a complex intervention tailored to real-world practice. Different methodological aspects of our study merit further discussion. We followed the Cochrane risk of bias tool for randomized trials in the critical evaluation of the method (185,193). Risk of bias in a RCT should be evaluated according to 5 domains: 1. bias arising from the randomization process, 2. bias due to deviations from intended interventions, 3. bias due to missing outcome data, 4. bias in the measurement in the outcome, and 5. bias in the selection of the reported result (185).

Bias arising from the randomization process

If the randomization is successfully accomplished, an influence of either known or unknown confounder factors is avoided (182,193). We generated the two comparison arms using simple randomization and a computer-generated random assignment code

calculated by the AmbuFlex-system (4). No differences in baseline characteristics between the arms were found, and hence, I believe that risk of bias related to the randomization process was low.

Bias due to deviations from intended interventions

The effect of assignment to the intervention was based on intention-to-treat (ITT) analyses as recommended, whereas the effect of adhering to the intervention was supported by per-protocol analyses (193). However, if there is non-adherence to the assigned intervention, an ITT analysis is expected to underestimate the intervention effect that would have been seen had all patients adhered to the intervention (193). In our study, non-adherence of patients to the intervention may have affected the results. The patients received detailed written information about the intervention after randomization and were informed to take action by themselves if they declined to participate. We cannot rule out the possibility that some of the allocated patients may not have fully understood how to react during the intervention. Potentially, this would have made us underestimate the effect of, e.g. self-management and satisfaction. One the other hand, outcome measures were perhaps less likely to have been affected, as the patients could call the outpatient clinic instead of using the intervention to seek advice. Further, it was difficult to define adherence, as some patients with stable disease would not have had a need of contact during follow-up. Blinding was not appropriate in our study as it was a pragmatic trial that aimed to evaluate the effect of an intervention in patients who were aware of their care (193). We used a pre-randomization design, which prevented allocation disappointment and attrition bias (137). Furthermore, I do not believe that lack of blinding in our study contributed to group differences in healthcare delivery performed by clinicians, but I cannot rule out bias related to the lack of blinding. Overall, some concerns are appropriate due to lack of blinding and non-adherence to the intervention.

Bias due to missing outcome data

The primary outcome including number of outpatient contacts was register based, and complete data were available for all participants in both groups, despite the seven patients who died during follow-up. Hence, risk of bias due to missing outcome data is related to a potential loss of self-reported secondary outcomes. Only approximately 60% of the patients responded to the follow-up research questionnaire in both groups. We have no information about the true value of patients with missing self-reported data, nor the process that led to data being missing. For example, patients with lower mental well-being were potentially less likely to respond to the follow-up questionnaire, and then the missing data of well-being depends on its true value. The missing well-being data will lead to bias if the missing data depend on both the true value and the assigned groups (193). It is not possible to prove whether the missingness in our study depended on both the true value and the assigned groups. We conducted sensitivity analysis to examine the

magnitude of missingness regarding mental well-being. We used "baseline observation carried forward" (193) and changed the baseline value in four analyses. The results indicated that the effect could either be overestimated or underestimated depending on the criteria. Multiple imputation is another approach to handle missing data (187); however, in this method, the missingness needs to be explained by measured variables. Overall, there is low risk of bias regarding the primary outcome, but a high risk of bias regarding the secondary, self-reported outcomes.

Bias in the measurement in the outcome

Measurement of primary outcome was register-based, and misclassification of procedure codes with respect to telephone consultations and outpatient visits cannot be ruled out. However, I believe that the potential misclassification is non-differential, as it was unrelated to the intervention assignments. Secondary outcomes were based on validated questionnaires. The patients responded to the baseline research questionnaire before randomization; thus, any misclassification was unrelated to the intervention assignments. The outcome assessor of the follow-up research questionnaire was the included patients because the secondary outcomes were based on self-reported information; thus, blinding was impossible. The assessment of self-reported outcomes could potentially be influenced by the knowledge of the intervention patients received during our study, and therefore, some concern regarding risk of bias in both directions should be considered. However, risk of bias with respect to the primary outcome was low.

Bias in the selection of the reported result

Bias in the selection of the reported results is considered low in our study. We devised a statistical analysis plan, and we followed this plan in the reporting. Additionally, we performed explorative analyses to investigate whether the intervention had an effect on subgroups in the population; however, results from these analyses were discussed as secondary findings.

Ethical considerations

We used a pre-randomization design in which group allocation information is only given to the participants in the intervention group after randomization. The design introduces some ethical issues that are not present in conventional RCTs, since the control arm participants are not informed about group assignment after randomization (137). Zelen points to three ethical issues: 1. should control group patients be informed about their assignment?, 2. should permission to use patient data be obtained?, and 3. is it proper to offer the intervention to only a group of patients? (137). According to the first issue, all patients in our study received a baseline and follow-up research questionnaire. However, patients in the control group were not informed about their assignment because they continued standard care. We could have included a question in the baseline questionnaire

regarding willingness to participate in a future RCT and made it known that only patients assigned to the intervention group would be contacted. This was done in "cohort multiple RCTs" from the Netherlands, whereby patients were asked for informed consent to be randomized in future RCTs (194). In that study, patients were informed that they would be offered the intervention if they were randomly selected; otherwise, they would serve as controls without being notified (194). If patients refused to participate in a future RCT, data from the cohort study could be used to provide information on the extent to which RCT participants represent the full cohort (195). With regard to the second issue, the patients in our study were informed that the questionnaire data were used for research. Other patient data were register-based, and approval from the Danish Data Protection Agency was obtained before enrollment of patients. Regarding the third issue, the intervention in our study was similar to what was offered to the control group. To test the effectiveness in a RCT design, only a subgroup of patients can be offered the intervention, as benefits and drawbacks have to be investigated before the intervention can be recommended for routine use. Despite these ethical considerations, I consider the pre-randomization to be feasible, since we had a large cohort of patients attending fixed-interval PRO-based follow-up, which was the target group for this study.

Selection of self-reported outcome measures

For PRO data to provide valuable and valid conclusions in RCTs, self-reported outcomes must be measured in a standardized manner using scales with robust measurement properties such as reliability, validity, and responsiveness (192). ISOQOL and COSMIN have both developed guidelines for selection of PRO measures for use in research (192,196). We selected several PRO instruments to measure secondary outcomes in our study, for example, the WHO-Five Well-Being (WHO-5) (87,88), the Health Literacy Questionnaire (HLQ) (107,108), and the General Self-Efficacy Scale (GSES) (109-111). These instruments aimed to evaluate the effect of the patient-initiated PRO-based intervention in relation to the patient perspective. We hypothesized that the level of well-being would be the same in the intervention and control arms and that improvements in health literacy and self-efficacy would be seen in the intervention arm. If PRO measures are used to evaluate whether an improvement has taken place in specific constructs in the intervention arm in a RCT, the instruments used to measure the change should be responsive. COSMIN defines responsiveness as "the ability of an instrument to detect change over time in the construct to be measured" (23). However, to my knowledge, the responsiveness of HLQ as an outcome in clinical trials has not been documented. In our study, the mean HLQ subscale 6 (ability to actively engage with healthcare providers) baseline scores were 3.87 in the intervention arm and 3.82 in the control arm, compared to 3.99 in the Danish population (108). In addition, to my knowledge, the responsiveness of GSES has not been reported. The GSES, however, has been used as an outcome in several

RCT studies. Some studies found an intervention effect on self-efficacy (197,198), while others did not (199-201). In our study, the mean GSES baseline scores were 29.35 in the intervention arm and 29.23 in the control arm. The mean GSES score is 32.87 in the Danish background population (111). Thus, in our study, the baseline scores of both HLQ and GSES were nearly the same as in the Danish background population. This led to a ceiling effect, and it was unlikely to measure improvement in the constructs during follow-up. In the design and planning phase, we did not identify studies that had used HLQ or GSES as an outcome measures in RCTs in an adult outpatient epilepsy population; nevertheless, we decided to use them. Generally, it was difficult to identify feasible instruments to measure these constructs and decide which to use.

GENERALIZABILITY

All studies took place in real-world clinical practice in a single outpatient clinic (Studies I and IV) or multiple outpatient clinics in the Central Denmark Region, with few inclusion criteria. Outpatient follow-up for patients with epilepsy is considered to be consistent in the Danish healthcare system; hence, I consider the results to be generalizable to other epilepsy outpatient populations.

However, the four studies consist of three different epilepsy outpatient populations (Figure 6 and Table 8). Study I included newly referred patients, Studies II and III included patients attending PRO-based follow-up (paper and web responders), and Study IV included patients attending PRO-based follow-up (web responders). Thus, Studies II, III, and IV represent selected patient populations, which is relevant to consider with regard to the generalizability of the results.

Lack of internal validity could affect the external validity. In Studies II and III, the reliability estimates are based on a selected homogeneous study population composed of a healthier group of patients than those in the target population. This could have underestimated the reliability, whereas the measurement error in Study III was affected to a lesser extent. In Study IV, non-adherence and loss of self-reported data at follow-up may have affected the study results in both directions. For these different reasons, caution must prevail with regard to the generalizability of the results in Studies II and IV.

Owing to the efforts to eliminate selection bias, Study I is considered to have high generalizability. In Studies I and III, I expect that similar results would be achieved in other epilepsy outpatient populations and perhaps also in other patient populations with chronic diseases.

CONCLUSION

This PhD project has contributed insights into different aspects of the use of remote PRO-based follow-up in outpatients with epilepsy. As of January 2020, remote PRO-based follow-up are used in outpatients with epilepsy at hospitals in all five Danish Regions. Based on the results and the discussion of the four studies in this PhD dissertation, the following conclusions can be drawn:

Among sociodemographic, personal, and disease-related variables, solo living, low education or household income, temporary or permanent social benefits, psychiatric diagnosis, female gender, low health literacy, self-efficacy, patient activation, well-being, or general health were associated with lower probability of referral to PRO-based follow-up, whereas age and co-morbidity were not noticeable factors. Overall, both register- and questionnaire-based data were consistent and indicated that socioeconomically advantaged patients were more likely to be referred to PRO-based follow-up than vulnerable patients. Further research should explore how healthcare services including PRO measures during outpatient follow-up can be of more support to less advantaged patients.

The PRO-based algorithm used to flag need for clinical attention in outpatient epilepsy clinics showed acceptable test-retest reliability. Different methods of administration produced similar results. However, lower reliability estimates were found if two different methods of administration were used. Test-retest of single questions in the epilepsy questionnaire showed fair to moderate reliability. Further evaluation of psychometric properties such as validity would be desirable in order to draw a more firm conclusion.

The Danish version of the WHO-Five Well-Being Index showed acceptable test-retest reliability, also across different methods of administration in an epilepsy outpatient population. However, the WHO-5 scale showed a relatively large measurement error, which should be taken into account when evaluating changes in well-being over time. Further research is required to explore the reliability of the WHO-5 scale in other language versions and patient populations.

Patient-initiated PRO-based follow-up did not show less use of healthcare resources or improved patient self-management or satisfaction compared to fixed-interval PRO-based follow-up. There is insufficient evidence for recommending a system-wide shift to patient-initiated PRO-based follow-up, but this model of care may be used as an

alternative to fixed-interval PRO-based follow-up in patients who prefer having an active role during their follow-up. Patients' individual self-management skills should be considered and careful introduction should be given before enrollment in a patient-initiated follow-up program. The effect of patient-initiated follow-up should be further investigated, and, preferably, patients with this type of follow-up should be compared to patients with scheduled in-clinic visits.

Overall, the findings of this dissertation have paid attention to reliability aspects and the diversity in the use of patient-level PRO measures in remote outpatient follow-up. Patient characteristics and preferences should be taken into consideration regarding both delivery and usage of healthcare services. The ambitions of using PRO measures in clinical practice for the benefit of vulnerable patients should be a future point of attention.

PERSPECTIVE AND FUTURE RESEARCH

Remote use of PRO measures in outpatient follow-up is a relatively new initiative in clinical practice that aims to enhance patient-centered care and flexible individual scheduling of hospital appointments. The large-scale implementation of AmbuFlex PRO solutions that has taken place since 2012 in the Danish healthcare system can provide material for further investigation. As of January 2020, AmbuFlex has been implemented in 35 different patient groups. Generally, there is a lack of evidence regarding all aspects related to the use of PRO measures in remote follow-up. This PhD project provides insight into some aspects related to use of PRO measures in remote outpatient follow-up. The results contribute with knowledge which could be of advantage regarding the implementation of this new model of care in a daily clinical setting in outpatients with epilepsy. However, there are still many perspectives that could benefit from further research.

Findings from our cohort study indicated that clinicians are aware of several aspects before referring a patient to PRO-based follow-up, as socioeconomically advantaged patients were more likely to be referred to PRO-based follow-up than vulnerable patients. This type of follow-up is not tailored to the whole patient population. Patients' capability and willingness to participate are of major importance, and the decision regarding referral should be based on a shared decision between the patient and the clinician that enables patients to freely choose their preferred method of outpatient follow-up. This could also prevent non-response and dropouts during follow-up. However, our findings also indicate that further research is needed. It could be beneficial to investigate how healthcare services can to a larger extent be supportive of vulnerable patients. This does not exclude use of PRO measures during follow-up, but may induce differential use of PRO measures, such as using PRO measures prior to telephone consultations or in-clinic visits, or as proxy solutions. For example, the Department of Neurology at Aarhus University Hospital has developed a PRO-based proxy version for epilepsy outpatients with cognitive disabilities, by which a relative or social worker completes the questionnaire on behalf of the patient. As of January 2020, 154 patients were attending the PRO-based proxy solution at two outpatient clinics in the Central Denmark Region. Furthermore, it could also be beneficial to investigate factors associated with patients who have dropped out of PRO-based follow-up. Qualitative research could be beneficial to

explore reasons related to participation refusal that are based on both patients' initiatives and clinicians' decisions.

The epilepsy questionnaire and the associated PRO-based algorithm described in our study are used in daily clinical practice for outpatients with epilepsy, and a large-scale implementation has taken place in the five Danish Regions since 2016. The questionnaire and algorithm have not been evaluated or revised since 2016 due to the national implementation process. The quality of the properties of the questionnaire measurements is of major importance. Results from the test-retest study should be discussed and evaluated by clinical experts, patients, and researchers during the next revision of the questionnaire. Future work should evaluate the validity of single items and the PRO-based algorithm. According to a legal change in the European Union's medical device regulation in 2018, the PRO-based algorithms used in AmbuFlex are classified as a medical device, and therefore, the PRO-based algorithms must be certified with a certification marking (75). In order to meet this demand, the PRO-based algorithms' validity has to be documented in the coming years.

The large measurement error found in the WHO-5 scale should also be discussed, as should the interpretation of the scale if clinicians are to use the scale to monitor change in mental well-being over time. The finding is also relevant to researchers who consider using the scale as an outcome measure in a clinical trial, as the measurement error found in our study is above the previously reported clinical relevant change of 10 points (88). Furthermore, it is highly relevant to investigate the test-retest reliability and measurement error of the WHO-5 scale in other patient populations and language versions. Furthermore, evaluation of psychometric properties such as responsiveness and what constitutes a minimal important change in the WHO-5 score is also desirable.

The healthcare system seeks new models of care to support more effective and patient-centered care. New models of care need to be evaluated in order to identify possible benefits or drawbacks. Use of PRO measures in remote follow-up often differs in scope and procedures between patient groups. We set out to investigate the effect of a new patient-initiated PRO-based intervention; however, the intervention did not provide evidence that supports a system-wide shift from fixed-interval to patient-initiated PRO-based follow-up in outpatients with epilepsy. If it is decided to use the patient-initiated model, referral should be based on a shared decision between the patient and clinician. The patient's self-management skills should be taken into consideration and the implementation process should also be planned and discussed carefully. Nevertheless, there remains a need for investigating whether a patient-initiated PRO-based model of care could benefit the healthcare system. Mejdahl et al. has contributed with important insight to the patient perspective in fixed-interval PRO-based follow-up (155), and she has also collected qualitative data of patients who received patient-initiated PRO-based

follow-up. Qualitative exploration of these data can hopefully address the question of why the expected differences were not observed in our study.

Preferably, future RCTs should compare the advantages of the patient-initiated PRO-based intervention by comparing results in patients who receive the intervention with patients who receive conventional follow-up with scheduled in-clinic appointments. It could also be beneficial to explore the effectiveness of the intervention in other patient populations with long-term illnesses. Furthermore, the effectiveness of fixed-interval PRO-based follow-up should also be investigated. Currently, two RCTs are in progresses that investigate the effect of remote fixed-interval PRO-based follow-up compared to fixed-interval in-clinic appointments in patients with chronic kidney disease (68) and type 1 diabetes (202). In addition, a RCT in patients with lung cancer explores the effect of remote PRO-based monitoring compared to fixed-interval in-clinic appointments (203).

The studies that make up this PhD dissertation focuses on three of the phases in PRO-based follow-up presented in Figure 3: referral, the questionnaire, and the PRO-based algorithm. However, the last phase, review of PRO responses by clinicians including PRO data presentation, interpretation, and clinical decision-making based on PRO data, has not been investigated. This should be an area of further investigation. In addition, a few RCTs related to use of PRO measures in remote patient management were identified (Table 2). A next step should be a comprehensive systematic literature search and review of the literature on this topic.

ENGLISH SUMMARY

Patient-reported outcome (PRO) measures are increasingly being used at the patient level in the healthcare system. In the generic Danish PRO system, AmbuFlex, PRO measures are used as the basis for remote outpatient follow-up, called PRO-based follow-up. In PRO-based follow-up, patients receive fixed-interval disease-specific questionnaires instead of in-clinic appointments during outpatient follow-up. PRO measures are captured at patients' homes and used to flag whether a patient needs further clinical attention based on a PRO-based algorithm. Several aspects are important to consider in remote PRO-based follow-up, for example, patient referral, the questionnaire and the PRO-based algorithm, and review of the responses by clinicians. However, there is scant research related to the use of remote PRO-based follow-up and little is known about the effectiveness of different PRO-based follow-up models.

This PhD project aims to provide insight into different aspects of the use of remote PRO-based follow-up in outpatients with epilepsy. The project is based on four studies with different focuses.

Study I was designed as a prospective cohort study that aimed to identify sociodemographic, personal, and disease-related determinants associated with referral to PRO-based follow-up. The study included 802 outpatients with epilepsy from the Department of Neurology at Aarhus University Hospital, Aarhus, Denmark. The study found that a number of factors were associated with referral to PRO-based follow-up. The study found that both self-reported and register-based analyses indicated that socioeconomically advantaged patients were more often referred to PRO-based follow-up than were vulnerable patients. Further research should explore how healthcare services including PRO measures during outpatient follow-up can be supportive of vulnerable patients.

Studies II & III were two test-retest reliability studies that aimed to evaluate the test-retest reliability of the PRO-based algorithm and the epilepsy questionnaire, including the Danish version of the WHO-Five Well-Being Index (WHO-5). The studies also evaluated whether web- or paper-based methods of administration influenced the results. The studies included 554 outpatients with epilepsy from three neurological departments in the Central Denmark Region. Study II found that the PRO-based algorithm showed acceptable test-retest reliability and that different methods of administration produced similar results; however, lower reliability estimates were found if more than one method was used. Study III showed acceptable test-retest reliability of the WHO-5 scale across

different methods of administration. Further, a relatively large measurement error in the scale was identified, and this should be taken into account when evaluating changes in well-being over time.

Study IV was designed as a pragmatic randomized controlled trial that aimed to provide insight into the effects of a patient-initiated PRO-based intervention on healthcare resource utilization, quality of care, and the patient perspective. The study included 593 outpatients with epilepsy from the Department of Neurology at Aarhus University Hospital, Aarhus, Denmark. The patients were randomized to either 1) patient-initiated PRO-based follow-up or 2) fixed-interval PRO-based follow-up. The study found no differences in use of healthcare resources, quality of care, patient self-management, or patient satisfaction between the arms. A system-wide shift to patient-initiated PRO-based follow-up is not recommended in outpatients with epilepsy, though this model of care may be used as an alternative to fixed-interval PRO-based follow-up in patients who wish to play an active role during their follow-up.

DANSK RESUMÉ

PRO-data (Patient Reported Outcome Data) anvendes i stigende grad på individniveau i sundhedsvæsenet. AmbuFlex er et generisk PRO-baseret system, hvor patientens egne oplysninger danner omdrejningspunkt for patientens ambulante opfølgning på hospitalet. Dette kaldes PRO-baseret opfølgning. I PRO-baseret opfølgning anvendes et sygdomsspecifikt spørgeskema, der sendes til patienten med regelmæssige intervaller, og dermed kan patienten undgå at komme ind til fastlagte besøg på hospitalet. Patienten besvarer spørgeskemaet hjemmefra, og baseret på en algoritme bruges svaret til at vurdere, om patienten har behov for kontakt. Flere aspekter er vigtige at overveje ved anvendelse af PRO-baseret opfølgning, eksempelvis visitationen af patienter, spørgeskemaet og den tilhørende PRO-baserede algoritme samt klinikernes vurdering af patientens besvarelse. Der er sparsom forskning om anvendelse af PRO-baseret opfølgning, og kun få studier har undersøgt effekten af forskellige PRO-baserede opfølgningsmodeller.

Det overordnede formål med ph.d.-projektet er at undersøge forskellige aspekter ved anvendelse af PRO-baseret opfølgning hos ambulante epilepsipatienter. Projektet er baseret på fire studier med forskelligt fokus.

Studie I er designet som et prospektivt kohortestudie, hvor formålet var at identificere socio-demografiske, personlige og sygdomsrelaterede faktorer associeret med visitation til PRO-baseret opfølgning. Der blev i alt inkluderet 802 ambulante epilepsipatienter fra Neurologisk Afdeling ved Aarhus Universitetshospital. Studiet viste, at en række faktorer var associeret med visitation til PRO-baseret opfølgning. Både selvrapporterede og registerbaserede analyser indikerede, at ressourcestærke patienter i højere grad blev visiteret til PRO-baseret opfølgning sammenlignet med patienter med færre ressourcer. Fremtidig forskning bør undersøge, hvordan PRO-data kan anvendes i den ambulante opfølgning af sårbare patienter.

Studierne II og III er designet som to test-retest reliabilitetsstudier, hvor formålet var at evaluere reliabiliteten af det anvendte epilepsispørgeskema og den tilhørende PRO-baserede algoritme samt den danske version af WHO-5 Trivselsindekset. Formålet var også at undersøge om patientens svarmetode (web eller papir) påvirkede resultaterne. Der blev i alt inkluderet 554 ambulante epilepsipatienter fra Region Midtjylland. Studie II viste acceptabel test-retest reliabilitet af den PRO-baserede algoritme, og de forskellige svarmetoder viste konsistente resultater, dog var reliabiliteten lavere, hvis der blev anvendt forskellige svarmetoder. Studie III viste acceptabel test-retest reliabilitet af WHO-

5 skalaen på tværs af forskellige svarmetoder. Der blev dog fundet en relativ stor målefejl ved skalaen, og dette skal tages i betragtning, hvis skalaen skal bruges til at måle ændringer i trivsel over tid.

Studie IV er designet som et randomiseret kontrolleret forsøg, hvor formålet var at evaluere effekten af en patientinitieret PRO-baseret intervention. Formålet var at sammenligne ressourceforbrug, behandlingskvalitet og patientperspektivet i to former for PRO-baseret ambulant opfølgning. Der blev i alt inkluderet 593 ambulante epilepsipatienter fra Neurologisk Afdeling ved Aarhus Universitetshospital. Patienterne blev randomiseret til enten: 1) patientinitieret PRO-baseret opfølgning (spørgeskema eller anden henvendelse til hospitalet på patientens initiativ) eller 2) PRO-baseret opfølgning med faste spørgeskemaintervaller. Studiet viste ingen forskel i ressourceforbrug, behandlingskvalitet, sygdomsrelateret egenomsorg eller tilfredshed mellem de to opfølgningsmodeller. Patientinitieret PRO-baseret opfølgning kan anvendes som et alternativ til patienter, der foretrækker at have en aktiv rolle i deres ambulante opfølgning. Der er dog utilstrækkelig evidens for at anbefale udbredelse af patientinitieret PRO-baseret opfølgning til ambulante epilepsipatienter.

REFERENCES

(1) US Department of Health and Human Services Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. 2009. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Accessed 09/17, 2019.

(2) Nancy E. Mayo. Dictionary of Quality of Life and Health Outcomes Measurement. First Edition. 2015. International Society for Quality of Life Research.

(3) Greenhalgh J, Dalkin S, Gooding K, Gibbons E, Wright J, Meads D, et.al. Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. Health Serv Deliv Res 2017;5(2).

(4) Schougaard LM, Larsen LP, Jessen A, Sidenius P, Dorflinger L, de Thurah A, et al. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Qual Life Res 2016 Mar;25(3):525-534.

(5) Whear R, Abdul-Rahman AK, Boddy K, Thompson-Coon J, Perry M, Stein K. The clinical effectiveness of patient initiated clinics for patients with chronic or recurrent conditions managed in secondary care: a systematic review. PLoS One 2013 Oct 7;8(10):e74774.

(6) Taneja A, Su'a B, Hill AG. Efficacy of patient-initiated follow-up clinics in secondary care: a systematic review. Intern Med J 2014 Dec;44(12a):1156-1160.

(7) Yach D, Hawkes C, Gould CL, Hofman KJ. The global burden of chronic diseases: overcoming impediments to prevention and control. JAMA 2004 Jun 2;291(21):2616-2622.

(8) Healthcare in Denmark. An overview. Danish Ministry of Health. 2017. Available at: https://www.healthcaredenmark.dk/media/ykedbhsl/healthcare-dk.pdf. Accessed 09/18, 2019.

(9) Hvidberg MF, Johnsen SP, Davidsen M, Ehlers L. A Nationwide Study of Prevalence Rates and Characteristics of 199 Chronic Conditions in Denmark. Pharmacoecon Open 2019 Jul 24.

(10) Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: translating evidence into action. Health Aff (Millwood) 2001 Nov-Dec;20(6):64-78.

(11) Richards T. Power to the people—via Paris. *BMJ*2017. Available at: https://blogs.bmj.com/bmj/2017/01/20/tessa-richards-power-to-the-people-via-paris/. Accessed 11/20, 2019.

(12) Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013 Jan 28;346:f167.

(13) Wu AW, Predmore ZS. Patient-Reported Outcomes: Toward Better Measurement of Patient-Centered Care in CKD. J Am Soc Nephrol 2019 Apr;30(4):523-525.

(14) Calvert M, Kyte D, Price G, Valderas JM, Hjollund NH. Maximising the impact of patient reported outcome assessment for patients and society. BMJ 2019 Jan 24;364:k5267.

(15) Hjollund NHI, Valderas JM, Kyte D, Calvert MJ. Health Data Processes: A Framework for Analyzing and Discussing Efficient Use and Reuse of Health Data With a Focus on Patient-Reported Outcome Measures. J Med Internet Res 2019 May 21;21(5):e12412.

(16) Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? Qual Life Res 2009 Feb;18(1):115-123.

(17) Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjaer E, Schougaard LM. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. Interact J Med Res 2014 Feb 11;3(1):e5.

(18) Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. Milbank Q 2014 Dec;92(4):754-775.

(19) Porter I, Goncalves-Bradley D, Ricci-Cabello I, Gibbons C, Gangannagaripalli J, Fitzpatrick R, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. J Comp Eff Res 2016 Aug;5(5):507-519.

(20) Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA 1995 Jan 4;273(1):59-65.

(21) Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. Qual Life Res 2008 Nov;17(9):1125-1135.

(22) de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: a practical guide. Cambridge University Press, 2011.

(23) Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010 Jul;63(7):737-745.

(24) Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010 May;19(4):539-549.

(25) The COSMIN Taxonomy of Measurement Properties. Available at: https://www.cosmin.nl/tools/cosmin-taxonomy-measurement-properties/. Accessed 09/18 2019.

(26) Jensen RE, Rothrock NE, DeWitt EM, Spiegel B, Tucker CA, Crane HM, et al. The role of technical advances in the adoption and integration of patient-reported outcomes in clinical care. Med Care 2015 Feb;53(2):153-159.

(27) Jones JB, Snyder CF, Wu AW. Issues in the design of Internet-based systems for collecting patient-reported outcomes. Qual Life Res 2007 Oct;16(8):1407-1417.

(28) Muehlhausen W, Doll H, Quadri N, Fordham B, O'Donohoe P, Dogar N, et al. Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013. Health Qual Life Outcomes 2015 Oct 7;13:167-015-0362-x.

(29) Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value Health 2009 Jun;12(4):419-429.

(30) Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. Value Health 2008 Mar-Apr;11(2):322-333.

(31) Rutherford C, Costa D, Mercieca-Bebber R, Rice H, Gabb L, King M. Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis. Qual Life Res 2016 Mar;25(3):559-574.

(32) Marcano Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev 2015 Jul 27;(7):MR000042. doi(7):MR000042.

(33) Lohr KN. Applications of health status assessment measures in clinical practice. Overview of the third conference on advances in health status assessment. Med Care 1992 May;30(5 Suppl):MS1-14.

(34) Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Qual Life Res 2012 Oct;21(8):1305-1314.

(35) Higginson IJ, Carr AJ. Measuring quality of life: Using quality of life measures in the clinical setting. BMJ 2001 May 26;322(7297):1297-1300.

(36) Santana MJ, Feeny D. Framework to assess the effects of using patient-reported outcome measures in chronic care management. Qual Life Res 2014 Jun;23(5):1505-1513.

(37) Donaldson MS. Taking PROs and patient-centered care seriously: incremental and disruptive ideas for incorporating PROs in oncology practice. Qual Life Res 2008 Dec;17(10):1323-1330.

(38) Basch E. Patient-Reported Outcomes - Harnessing Patients' Voices to Improve Clinical Care. N Engl J Med 2017 Jan 12;376(2):105-108.

(39) Ishaque S, Karnon J, Chen G, Nair R, Salter AB. A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). Qual Life Res 2019 Mar;28(3):567-592.

(40) Fonhus MS, Dalsbo TK, Johansen M, Fretheim A, Skirbekk H, Flottorp SA. Patient-mediated interventions to improve professional practice. Cochrane Database Syst Rev 2018 Sep 11;9:CD012472.

(41) Kotronoulas G, Kearney N, Maguire R, Harrow A, Di Domenico D, Croy S, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. J Clin Oncol 2014 May 10;32(14):1480-1501.

(42) Boyce MB, Browne JP. Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? A systematic review. Qual Life Res 2013 Nov;22(9):2265-2278.

(43) Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. BMC Health Serv Res 2013 Jun 11;13:211-6963-13-211.

(44) Johansen MA, Berntsen GK, Schuster T, Henriksen E, Horsch A. Electronic symptom reporting between patient and provider for improved health care service quality: a systematic review of randomized controlled trials. part 2: methodological quality and effects. J Med Internet Res 2012 Oct 3;14(5):e126.

(45) Luckett T, Butow PN, King MT. Improving patient outcomes through the routine use of patient-reported data in cancer clinics: future directions. Psychooncology 2009 Nov;18(11):1129-1138.

(46) Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res 2008 Mar;17(2):179-193.

(47) Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J Eval Clin Pract 2006 Oct;12(5):559-568.

(48) Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals: a systematic review of its impact. Med Care 2000 Feb;38(2):175-186.

(49) Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. J Eval Clin Pract 1999 Nov;5(4):401-416.

(50) Yang LY, Manhas DS, Howard AF, Olson RA. Patient-reported outcome use in oncology: a systematic review of the impact on patient-clinician communication. Support Care Cancer 2018 Jan;26(1):41-60.

(51) 26th Annual Conference of the International Society for Quality of Life Research. Qual Life Res 2019 Oct;28(Suppl 1):1-190.

(52) de Thurah A, Stengaard-Pedersen K, Axelsen M, Fredberg U, Schougaard LMV, Hjollund NHI, et al. Tele-Health Followup Strategy for Tight Control of Disease Activity in Rheumatoid

Arthritis: Results of a Randomized Controlled Trial. Arthritis Care Res (Hoboken) 2018 Mar;70(3):353-360.

(53) de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. Lancet 2017 Sep 2;390(10098):959-968.

(54) Armstrong KA, Coyte PC, Brown M, Beber B, Semple JL. Effect of Home Monitoring via Mobile App on the Number of In-Person Visits Following Ambulatory Surgery: A Randomized Clinical Trial. JAMA Surg 2017 Jul 1;152(7):622-627.

(55) El Miedany Y, El Gaafary M, Youssef S, Bahlas S, Almedany S, Ahmed I, et al. Toward Electronic Health Recording: Evaluation of Electronic Patient-reported Outcome Measures System for Remote Monitoring of Early Rheumatoid Arthritis. J Rheumatol 2016 Dec;43(12):2106-2112.

(56) Denis F, Lethrosne C, Pourel N, Molinier O, Pointreau Y, Domont J, et al. Randomized Trial Comparing a Web-Mediated Follow-up With Routine Surveillance in Lung Cancer Patients. J Natl Cancer Inst 2017 Sep 1;109(9):10.1093/jnci/djx029.

(57) Denis F, Basch E, Septans AL, Bennouna J, Urban T, Dueck AC, et al. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. JAMA 2019 Jan 22;321(3):306-307.

(58) Lizee T, Basch E, Tremolieres P, Voog E, Domont J, Peyraga G, et al. Cost-effectiveness of web-based patient-reported outcome surveillance in lung cancer patients. J Thorac Oncol 2019 Feb 15.

(59) Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. J Clin Oncol 2016 Feb 20;34(6):557-565.

(60) Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. JAMA 2017 Jul 11;318(2):197-198.

(61) Wheelock AE, Bock MA, Martin EL, Hwang J, Ernest ML, Rugo HS, et al. SIS.NET: a randomized controlled trial evaluating a web-based system for symptom management after treatment of breast cancer. Cancer 2015 Mar 15;121(6):893-899.

(62) Mooney KH, Beck SL, Wong B, Dunson W, Wujcik D, Whisenant M, et al. Automated home monitoring and management of patient-reported symptoms during chemotherapy: results of the symptom care at home RCT. Cancer Med 2017 Mar;6(3):537-546.

(63) Mooney K, Whisenant MS, Beck SL. Symptom Care at Home: A Comprehensive and Pragmatic PRO System Approach to Improve Cancer Symptom Care. Med Care 2019 May;57 Suppl 5 Suppl 1:S66-S72.

(64) Kyte D, Bishop J, Brettell E, Calvert M, Cockwell P, Dutton M, et al. Use of an electronic patient-reported outcome measure in the management of patients with advanced chronic kidney disease: the RePROM pilot trial protocol. BMJ Open 2018 Oct 28;8(10):e026080-2018-026080.

(65) Girgis A, Durcinoska I, Gerges M, Kaadan N, Arnold A, Descallar J, et al. Study protocol for a controlled trial of an eHealth system utilising patient reported outcome measures for personalised treatment and care: PROMPT-Care 2.0. BMC Cancer 2018 Aug 23;18(1):845-018-4729-3.

(66) Maguire R, Fox PA, McCann L, Miaskowski C, Kotronoulas G, Miller M, et al. The eSMART study protocol: a randomised controlled trial to evaluate electronic symptom management using the advanced symptom management system (ASyMS) remote technology for patients with cancer. BMJ Open 2017 Jun 6;7(5):e015016-2016-015016.

(67) Warrington L, Absolom K, Holch P, Gibson A, Clayton B, Velikova G. Online tool for monitoring adverse events in patients with cancer during treatment (eRAPID): field testing in a clinical setting. BMJ Open 2019 Feb 1;9(1):e025185-2018-025185.

(68) Grove BE, Ivarsen P, de Thurah A, Schougaard LM, Kyte D, Hjollund NH. Remote follow-up using patient-reported outcome measures in patients with chronic kidney disease: the PROKID study - study protocol for a non-inferiority pragmatic randomised controlled trial. BMC Health Serv Res 2019 Sep 4;19(1):631-019-4461-y.

(69) Kumarakulasingam P, McDermott H, Patel N, Boutler L, Tincello DG, Peel D, et al. Acceptability and utilisation of patient-initiated follow-up for endometrial cancer amongst women from diverse ethnic and social backgrounds: A mixed methods study. Eur J Cancer Care (Engl) 2019 Mar;28(2):e12997.

(70) Jeppesen MM, Jensen PT, Hansen DG, Christensen RD, Mogensen O. Patient-initiated follow up affects fear of recurrence and healthcare use: a randomised trial in early-stage endometrial cancer. BJOG 2018 Jul 6.

(71) Batehup L, Porter K, Gage H, Williams P, Simmonds P, Lowson E, et al. Follow-up after curative treatment for colorectal cancer: longitudinal evaluation of patient initiated follow-up in the first 12 months. Support Care Cancer 2017 Jul;25(7):2063-2073.

(72) Kirshbaum MN, Dent J, Stephenson J, Topping AE, Allinson V, McCoy M, et al. Open access follow-up care for early breast cancer: a randomised controlled quality of life analysis. Eur J Cancer Care (Engl) 2017 Jul;26(4):10.1111/ecc.12577. Epub 2016 Sep 26.

(73) Whear R, Abdul-Rahman AK, Thompson-Coon J, Boddy K, Perry MG, Stein K. Patient initiated clinics for patients with chronic or recurrent conditions managed in secondary care: a systematic review of patient reported outcomes and patient and clinician satisfaction. BMC Health Serv Res 2013 Dec 1;13:501-6963-13-501.

(74) Langstrup H. Patient-reported data and the politics of meaningful data work. Health Informatics J 2018 Dec 31:1460458218820188.

(75) Hjollund NHI. Fifteen Years' Use of Patient-Reported Outcome Measures at the Group and Patient Levels: Trend Analysis. J Med Internet Res 2019 Sep 30;21(9):e15856.

(76) Christensen J, FAU - Vestergaard M, Vestergaard M, FAU - Pedersen MG, Pedersen MG, FAU - Pedersen CB, et al. Incidence and prevalence of epilepsy in Denmark. - Epilepsy Res.2007 Aug;76(1):60-5.Epub 2007 Aug 7. (0920-1211 (Print); 0920-1211 (Linking).

(77) Fiest KM, Sauro KM, Wiebe S, Patten SB, Kwon CS, Dykeman J, et al. Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies. Neurology 2017 Jan 17;88(3):296-303.

(78) Fisher RS, van Emde Boas W, Blume W, Elger C, Genton P, Lee P, et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia 2005 Apr;46(4):470-472.

(79) Josephson CB, Patten SB, Bulloch A, Williams JVA, Lavorato D, Fiest KM, et al. The impact of seizures on epilepsy outcomes: A national, community-based survey. Epilepsia 2017 May;58(5):764-771.

(80) Nixon A, Kerr C, Breheny K, Wild D. Patient Reported Outcome (PRO) assessment in epilepsy: a review of epilepsy-specific PROs according to the Food and Drug Administration (FDA) regulatory requirements. Health Qual Life Outcomes 2013 Mar 11;11:38-7525-11-38.

(81) A structured review of patient-reported outcome measures (PROMs) for epilepsy. 2009. Patient-reported outcome measurement group, Oxford, UK. Available at: https://phi.uhce.ox.ac.uk/pdf/PROMs_Oxford_Epilepsy_17092010.pdf. Accessed 12/04, 2019.

(82) Stafford M, Gavriel S, Lloyd A. Patient-reported outcomes measurements in epilepsy. Expert Rev Pharmacoecon Outcomes Res 2007 Aug;7(4):373-384.

(83) Jones FJS, Ezzeddine FL, Herman ST, Buchhalter J, Fureman B, Moura LMVR. A feasibility assessment of functioning and quality-of-life patient-reported outcome measures in adult epilepsy clinics: A systematic review. Epilepsy Behav 2020 Jan;102:106704.

(84) Wagner AK, Vickrey BG. The routine use of health-related quality of life measures in the care of patients with epilepsy: rationale and research agenda. Qual Life Res 1995 Apr;4(2):169-177.

(85) Moura LM, Schwamm E, Moura Junior V, Seitz MP, Hsu J, Cole AJ, et al. Feasibility of the collection of patient-reported outcomes in an ambulatory neurology clinic. Neurology 2016 Dec 6;87(23):2435-2442.

(86) Bergmann M, Prieschl M, Walser G, Luef G, Rumpold G, Unterberger I. Computer-based monitoring and evaluation of epilepsy-related health variables and their impact on treatment decision. Epilepsy Behav 2018 Jul;84:173-178.

(87) Staehr Johansen K: The use of well-being measures in primary health care – the Dep-Care project; in World Health Organization, Regional Office for Europe: Well-Being Measures in Primary Health Care – the DepCare Project. Geneva, World Health Organization, 1998, target 12, E60246.

(88) Topp CW, Ostergaard SD, Sondergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176.

(89) Bjorner JB, Thunedborg K, Kristensen TS, Modvig J, Bech P. The Danish SF-36 Health Survey: translation and preliminary validity studies. J Clin Epidemiol 1998 Nov;51(11):991-999.

(90) Bjorner JB, Damsgaard MT, Watt T, Groenvold M. Tests of data quality, scaling assumptions, and reliability of the Danish SF-36. J Clin Epidemiol 1998 Nov;51(11):1001-1011.

(91) Olsen LR, Mortensen EL, Bech P. The SCL-90 and SCL-90R versions validated by item response models in a Danish community sample. Acta Psychiatr Scand 2004 Sep;110(3):225-229.

(92) Bronserud MM, Iachina M, Green A, Groenvold M, Dorflinger L, Jakobsen E. Patient-reported outcomes (PROs) in lung cancer: Experiences from a nationwide feasibility study. Lung Cancer 2019 Feb;128:67-73.

(93) Multone E, Vader JP, Mottet C, Schoepfer A, Fournier N, Burnand B, et al. Characteristics of non-responders to self-reported questionnaires in a large inflammatory bowel disease cohort study. Scand J Gastroenterol 2015;50(11):1348-1356.

(94) Christensen AI, Ekholm O, Glumer C, Andreasen AH, Hvidberg MF, Kristensen PL, et al. The Danish National Health Survey 2010. Study design and respondent characteristics. Scand J Public Health 2012 Jun;40(4):391-397.

(95) Abrahamsen R, Svendsen MV, Henneberger PK, Gundersen GF, Toren K, Kongerud J, et al. Non-response in a cross-sectional study of respiratory health in Norway. BMJ Open 2016 Jan 6;6(1):e009912-2015-009912.

(96) Kjaer ASHK, Rasmussen TA, Hjollund NH, Rodkjaer LO, Storgaard M. Patient-reported outcomes in daily clinical practise in HIV outpatient care. Int J Infect Dis 2018 Apr;69:108-114.

(97) Jamilloux Y, Sarabi M, Kerever S, Boussely N, le Sidaner A, Valgueblasse V, et al. Adherence to online monitoring of patient-reported outcomes by patients with chronic inflammatory diseases: a feasibility study. Lupus 2015 Nov;24(13):1429-1436.

(98) Christensen AI, Ekholm O, Gray L, Glumer C, Juel K. What is wrong with non-respondents? Alcohol-, drug- and smoking-related mortality and morbidity in a 12-year follow-up study of respondents and non-respondents in the Danish Health and Morbidity Survey. Addiction 2015 Sep;110(9):1505-1512.

(99) Halliday JA, Hendrieckx C, Busija L, Browne JL, Nefs G, Pouwer F, et al. Validation of the WHO-5 as a first-step screening instrument for depression in adults with diabetes: Results from Diabetes MILES - Australia. Diabetes Res Clin Pract 2017 Oct;132:27-35.

(100) Hajos TR, Pouwer F, Skovlund SE, Den Oudsten BL, Geelhoed-Duijvestijn PH, Tack CJ, et al. Psychometric and screening properties of the WHO-5 well-being index in adult outpatients with Type 1 or Type 2 diabetes mellitus. Diabet Med 2013 Feb;30(2):e63-9.

(101) Guethmundsdottir HB, Olason DP, Guethmundsdottir DG, Sigurethsson JF. A psychometric evaluation of the Icelandic version of the WHO-5. Scand J Psychol 2014 Dec;55(6):567-572.

(102) Krieger T, Zimmermann J, Huffziger S, Ubl B, Diener C, Kuehner C, et al. Measuring depression with a well-being index: further evidence for the validity of the WHO Well-Being Index (WHO-5) as a measure of the severity of depression. J Affect Disord 2014 Mar;156:240-244.

(103) de Wit M, Pouwer F, Gemke RJ, Delemarre-van de Waal HA, Snoek FJ. Validation of the WHO-5 Well-Being Index in adolescents with type 1 diabetes. Diabetes Care 2007 Aug;30(8):2003-2006.

(104) Sundheds- og Ældreministeriet. Bekendtgørelse af lov om videnskabsetisk behandling af sundhedsvidenskabelige forskningsprojekter [In Danish]. 2017; Available at: https://www.retsinformation.dk/Forms/r0710.aspx?id=192671#idfae75201-60d1-46a6-bef7-ccd4ce8f39aa. Accessed 09/23, 2019.

(105) Datatilsynet. Vejledning Samtykke [In Danish]. 2019; Available at: https://www.datatilsynet.dk/media/6562/samtykke.pdf. Accessed 09/23, 2019.

(106) Business Intelligence in Central Denmark Region [In Danish]. Available at: https://www.rm.dk/om-os/organisation/it/business-intelligence/. Accessed 09/27, 2019.

(107) Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013 Jul 16;13:658-2458-13-658.

(108) Maindal HT, Kayser L, Norgaard O, Bo A, Elsworth GR, Osborne RH. Cultural adaptation and validation of the Health Literacy Questionnaire (HLQ): robust nine-dimension Danish language confirmatory factor model. Springerplus 2016 Aug 2;5(1):1232-016-2887-9. eCollection 2016.

(109) Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol 2005 Sep;139(5):439-457.

(110) Schwarzer R, & Jerusalem M. Generalized Self-Efficacy scale. In J. Weinman, S.Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs.1995, (pp. 35-37). Windsor, England: NFER-NELSON.

(111) Scholz U, Gutiérrez-Doña B, Sud S, Schwarzer R. Is general self-efficacy a universal construct? Psychometric findings from 25 countries. European Journal of Psychological Assessment, 2002, 18 (3), 242-251.

(112) Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form Patient Activation Measure (PAM13) in a Danish version. BMC Public Health 2009 Jun 29;9:209-2458-9-209.

(113) Danmarks Statistik. Brug af forskermaskiner [In Danish]. Available at: https://www.dst.dk/da/TilSalg/Forskningsservice/brug-af-forskermaskiner. Accessed 11/05, 2019.

(114) Pedersen CB. The Danish Civil Registration System. Scand J Public Health 2011 Jul;39(7 Suppl):22-25.

(115) Frank L. Epidemiology. When an entire country is a cohort. Science 2000 Mar 31;287(5462):2398-2399.

(116) Jensen VM, Rasmussen AW. Danish Education Registers. Scand J Public Health 2011 Jul;39(7 Suppl):91-94.

(117) Baadsgaard M, Quitzau J. Danish registers on personal income and transfer payments. Scand J Public Health 2011 Jul;39(7 Suppl):103-105.

(118) Hjollund NH, Larsen FB, Andersen JH. Register-based follow-up of social benefits and other transfer payments: accuracy and degree of completeness in a Danish interdepartmental administrative database compared with a population-based survey. Scand J Public Health 2007;35(5):497-502.

(119) Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health 2011 Jul;39(7 Suppl):30-33.

(120) Schougaard LMV, de Thurah A, Christensen J, Lomborg K, Maindal HT, Mejdahl CT, et al. Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study. Qual Life Res 2020 Jan 3.

(121) UNESCO. International Standard Classification of Education. 1998; Available at: http://uis.unesco.org/en/topic/international-standard-classification-education-isced. Accessed 09/06, 2019.

(122) Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383.

(123) The Psychiatric Research Unit at the Mental Health Centre North Zealand, Denmark. The WHO-Five Well-being Index. Available at: www.who-5.org. Accessed 12/09, 2019.

(124) Overgaard MK, Andersen PK, Parner ET. Regression analysis of censored data using pseudo-observations: An update. Stata Journal 2015; 15 (3): 809-821.

(125) Klein JP, Logan B, Harhoff M, Andersen PK. Analyzing survival curves at a fixed point in time. Stat Med 2007 Oct 30;26(24):4505-4519.

(126) Biering K, Hjollund NH, Frydenberg M. Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. Clin Epidemiol 2015 Jan 16;7:91-106.

(127) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009 Jun 29;338:b2393.

(128) Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. Value Health 2011 Dec;14(8):967-977.

(129) Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. J Adv Nurs 2003 Apr;42(1):57-63.

(130) COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 2019; Available at: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf#. Accessed 09/17, 2019.

(131) Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011 Jan;64(1):96-106.

(132) Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174.

(133) Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016 Jun;15(2):155-163.

(134) Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007 Jan;60(1):34-42.

(135) Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol 2010 Aug;63(8):834-840.

(136) Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. JAMA 2013 Feb 27;309(8):814-822.

(137) Zelen M. A new design for randomized clinical trials. N Engl J Med 1979 May 31;300(22):1242-1245.

(138) Schougaard LM, Mejdahl CT, Petersen KH, Jessen A, de Thurah A, Sidenius P, et al. Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: study protocol for a pragmatic randomised controlled study. BMC Health Serv Res 2017 Jan 26;17(1):83-017-2015-8.

(139) Spørgeskema. En undersøgelse af dine behov og oplevelser under udredning og behandling for kræft [In Danish]. Danish Cancer Society. 2017. Available at: https://www.cancer.dk/dyn/resources/File/file/2/6912/1513860458/final_barometerskema_nyeste.pdf. Accessed 11/14, 2018.

(140) Williams JG, Cheung WY, Russell IT, Cohen DR, Longo M, Lervy B. Open access follow up for inflammatory bowel disease: pragmatic randomised trial and cost effectiveness study. BMJ 2000 Feb 26;320(7234):544-548.

(141) Stata Corporation, College Station, Texas. Bootstrap manual. Available at: https://www.stata.com/manuals13/rbootstrap.pdf. Accessed 11/27, 2019.

(142) Schougaard LMV, de Thurah A, Christiansen DH, Sidenius P, Hjollund NH. Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test-retest reliability study. BMJ Open 2018 Jul 25;8(7):e021337-2017-021337.

(143) Schougaard LMV, de Thurah A, Bech P, Hjollund NH, Christiansen DH. Test-retest reliability and measurement error of the Danish WHO-5 Well-being Index in outpatients with epilepsy. Health Qual Life Outcomes 2018 Sep 6;16(1):175-018-1001-0.

(144) Schougaard LMV, Mejdahl CT, Christensen J, Lomborg K, Maindal HT, de Thurah A, et al. Patient-initiated versus fixed-interval patient-reported outcome-based follow-up in outpatients with epilepsy: a pragmatic randomized controlled trial. J Patient Rep Outcomes 2019 Sep 13;3(1):61-019-0151-0.

(145) Hayden JA, Cote P, Steenstra IA, Bombardier C, QUIPS-LBP Working Group. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. J Clin Epidemiol 2008 Jun;61(6):552-560.

(146) Hutchings A, Neuburger J, Grosse Frie K, Black N, van der Meulen J. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. Health Qual Life Outcomes 2012 Mar 30;10:34-7525-10-34.

(147) Bobitt J, Aguayo L, Payne L, Jansen T, Schwingel A. Geographic and Social Factors Associated With Chronic Disease Self-Management Program Participation: Going the "Extra-Mile" for Disease Prevention. Prev Chronic Dis 2019 Mar 07;16:10.5888/pcd16.180385.

(148) Verevkina N, Shi Y, Fuentes-Caceres VA, Scanlon DP. Attrition in Chronic Disease Self-Management Programs and self-efficacy at enrollment. Health Educ Behav 2014 Dec;41(6):590-598.

(149) Tecson KM, Bass K, Felius J, Hall SA, Jamil AK, Carey SA. Patient "Activation" of Patients Referred for Advanced Heart Failure Therapy. Am J Cardiol 2019 Feb 15;123(4):627-631.

(150) Jeppesen MM, Mogensen O, Hansen DG, Iachina M, Korsholm M, Jensen PT. Detection of recurrence in early stage endometrial cancer - the role of symptoms and routine follow-up. Acta Oncol 2017 Feb;56(2):262-269.

(151) Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002 Oct -Nov;48(2):177-187.

(152) Abu Dabrh AM, Gallacher K, Boehmer KR, Hargraves IG, Mair FS. Minimally disruptive medicine: the evidence and conceptual progress supporting a new era of healthcare. J R Coll Physicians Edinb 2015;45(2):114-117.

(153) Mejdahl CT, Schougaard LMV, Hjollund NH, Riiskjaer E, Lomborg K. Exploring organisational mechanisms in PRO-based follow-up in routine outpatient care - an interpretive description of the clinician perspective. BMC Health Serv Res 2018 Jul 13;18(1):546-018-3352-y.

(154) Knudsen LR, de Thurah A, Lomborg K. Experiences With Telehealth Followup in Patients With Rheumatoid Arthritis: A Qualitative Interview Study. Arthritis Care Res (Hoboken) 2018 Sep;70(9):1366-1372.

(155) Mejdahl CT, Schougaard LMV, Hjollund NH, Riiskjaer E, Thorne S, Lomborg K. PRO-based follow-up as a means of self-management support - an interpretive description of the patient perspective. J Patient Rep Outcomes 2018 Sep 5;2:38-018-0067-0. eCollection 2017.

(156) Philpot LM, Barnes SA, Brown RM, Austin JA, James CS, Stanford RH, et al. Barriers and Benefits to the Use of Patient-Reported Outcome Measures in Routine Clinical Care: A Qualitative Study. Am J Med Qual 2018 Jul;33(4):359-364.

(157) Holmes MM, Stanescu S, Bishop FL. The Use of Measurement Systems to Support Patient Self-Management of Long-Term Conditions: An Overview of Opportunities and Challenges. Patient Relat Outcome Meas 2019 Dec 16;10:385-394.

(158) Mejdahl, C., Nielsen, B., Hjøllund, N. H., Lomborg, K. Use of patientreported outcomes in outpatient settings as a means of patient involvement and self-management support - a qualitative study of the patient perspective. Eur J Pers Cent Healthc, 2016; 4: 359–367.

(159) Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005 May;37(5):360-363.

(160) It-anvendelse i befolkningen. Danmarks Statistik. 2016. [In Danish]. Available at: https://www.dst.dk/Site/Dst/Udgivelser/GetPubFile.aspx?id=20738&sid=itbef2016. Accessed 12/16, 2019.

(161) Rasmussen SL, Rejnmark L, Ebbehoj E, Feldt-Rasmussen U, Rasmussen AK, Bjorner JB, et al. High Level of Agreement between Electronic and Paper Mode of Administration of a Thyroid-Specific Patient-Reported Outcome, ThyPRO. Eur Thyroid J 2016 Mar;5(1):65-72.

(162) Egger MJ, Lukacz ES, Newhouse M, Wang J, Nygaard I. Web versus paper-based completion of the epidemiology of prolapse and incontinence questionnaire. Female Pelvic Med Reconstr Surg 2013 Jan-Feb;19(1):17-22.

(163) Sjostrom M, Stenlund H, Johansson S, Umefjord G, Samuelsson E. Stress urinary incontinence and quality of life: a reliability study of a condition-specific instrument in paper and web-based versions. Neurourol Urodyn 2012 Nov;31(8):1242-1246.

(164) Salaffi F, Gasparini S, Ciapetti A, Gutierrez M, Grassi W. Usability of an innovative and interactive electronic system for collection of patient-reported data in axial spondyloarthritis: comparison with the traditional paper-administered format. Rheumatology (Oxford) 2013 Nov;52(11):2062-2070.

(165) Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res 2018 May;27(5):1159-1170.

(166) Lucas-Carrasco R. Reliability and validity of the Spanish version of the World Health Organization-Five Well-Being Index in elderly. Psychiatry Clin Neurosci 2012 Oct;66(6):508-513.

(167) Newnham EA, Hooke GR, Page AC. Monitoring treatment response and outcomes using the World Health Organization's Wellbeing Index in psychiatric care. J Affect Disord 2010 Apr;122(1-2):133-138.

(168) van Dijk SEM, Adriaanse MC, van der Zwaan L, Bosmans JE, van Marwijk HWJ, van Tulder MW, et al. Measurement properties of depression questionnaires in patients with diabetes: a systematic review. Qual Life Res 2018 Jun;27(6):1415-1430.

(169) Dadfar M, Momeni Safarabad N, Asgharnejad Farid AA, Nemati Shirzy M, Ghazie Pour Abarghouie F. Reliability, validity, and factorial structure of the World Health Organization-5 Well-Being Index (WHO-5) in Iranian psychiatric outpatients. Trends Psychiatry Psychother 2018;40(2):79-84.

(170) Garland AF, Deyessa N, Desta M, Alem A, Zerihun T, Hall KG, et al. Use of the WHO's Perceived Well-Being Index (WHO-5) as an efficient and potentially valid screen for depression in a low income country. Fam Syst Health 2018 Jun;36(2):148-158.

(171) Bonnin CM, Yatham LN, Michalak EE, Martinez-Aran A, Dhanoa T, Torres I, et al. Psychometric properties of the well-being index (WHO-5) spanish version in a sample of euthymic patients with bipolar disorder. J Affect Disord 2018 Mar 1;228:153-159.

(172) Englbrecht M, Alten R, Aringer M, Baerwald CG, Burkhardt H, Eby N, et al. Validation of Standardized Questionnaires Evaluating Symptoms of Depression in Rheumatoid Arthritis Patients: Approaches to Screening for a Frequent Yet Underrated Challenge. Arthritis Care Res (Hoboken) 2017 Jan;69(1):58-66.

(173) Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD, Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health 2007 Nov-Dec;10 Suppl 2:S94-S105.

(174) Hewlett S, Kirwan J, Pollock J, Mitchell K, Hehir M, Blair PS, et al. Patient initiated outpatient follow up in rheumatoid arthritis: six year randomised controlled trial. BMJ 2005 Jan 22;330(7484):171.

(175) Robinson A, Thompson DG, Wilkin D, Roberts C, Northwest Gastrointestinal Research Group. Guided self-management and patient-directed follow-up of ulcerative colitis: a randomised trial. Lancet 2001 Sep 22;358(9286):976-981.

(176) John K, Tailor S, Anderson J, Lawthom C. Managing epilepsy in austerity - Evaluating the utility and value of the epilepsy specialist nurse in an open access model of service delivery. Aneurin Bevan Epilepsy Specialist Team (A.B.E.S.T.). Seizure 2019 Feb;65:98-100.

(177) Berkhof FF, Hesselink AM, Vaessen DL, Uil SM, Kerstjens HA, van den Berg JW. The effect of an outpatient care on-demand-system on health status and costs in patients with COPD. A randomized trial. Respir Med 2014 Aug;108(8):1163-1170.

(178) Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152.

(179) Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017 Nov 1;19(11):e367.

(180) May C, Montori VM, Mair FS. We need minimally disruptive medicine. BMJ 2009 Aug 11;339:b2803.

(181) Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655.

(182) Gordis. Epidemiology. Fifth Edition. Philadelphia, USA: Elsevier Saunders; 2014.

(183) Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med 2013 Feb 19;158(4):280-286.

(184) Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res 2018 May;27(5):1171-1179.

(185) Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898.

(186) Sorensen HT, Lash TL, Rothman KJ. Beyond randomized controlled trials: a critical comparison of trials with nonrandomized studies. Hepatology 2006 Nov;44(5):1075-1082.

(187) Rombach I, Gray AM, Jenkinson C, Murray DW, Rivero-Arias O. Multiple imputation for patient reported outcome measures in randomised controlled trials: advantages and disadvantages of imputing at the item, subscale or composite score level. BMC Med Res Methodol 2018 Aug 28;18(1):87-018-0542-6.

(188) Olsen J. Register-based research: some methodological considerations. Scand J Public Health 2011 May;39(3):225-229.

(189) Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sorensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol 2015 Nov 17;7:449-490.

(190) Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc 2016 May 04;9:211-217.

(191) COSMIN Risk of Bias checklist. 2018; Available at: https://www.cosmin.nl/wp-content/uploads/COSMIN-RoB-checklist-V2-0-v17_rev3.pdf. Accessed 12/12, 2019.

(192) Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res 2013 Oct;22(8):1889-1905.

(193) Revised Cochrane risk-of-bias tool for randomized trials (RoB 2). 2019. Available at: https://drive.google.com/file/d/19R9savfPdCHC8XLz2iiMvL_71lPJERWK/view. Accessed 12/07, 2019.

(194) van der Velden JM, Verkooijen HM, Young-Afat DA, Burbach JP, van Vulpen M, Relton C, et al. The cohort multiple randomized controlled trial design: a valid and efficient alternative to pragmatic trials? Int J Epidemiol 2017 Feb 1;46(1):96-102.

(195) Young-Afat DA, Verkooijen HA, van Gils CH, van der Velden JM, Burbach JP, Elias SG, et al. Brief Report: Staged-informed Consent in the Cohort Multiple Randomized Controlled Trial Design. Epidemiology 2016 May;27(3):389-392.

(196) COSMIN. Find the right tool. Available at: https://www.cosmin.nl/finding-right-tool/select-best-measurement-instrument/. Accessed 12/17, 2019.

(197) Haller N, Lorenz S, Pfirrmann D, Koch C, Lieb K, Dettweiler U, et al. Individualized Web-Based Exercise for the Treatment of Depression: Randomized Controlled Trial. JMIR Ment Health 2018 Oct 12;5(4):e10698.

(198) Mahlke CI, Priebe S, Heumann K, Daubmann A, Wegscheider K, Bock T. Effectiveness of one-to-one peer support for patients with severe mental illness - a randomised controlled trial. Eur Psychiatry 2017 May;42:103-110.

(199) Bringsvor HB, Langeland E, Oftedal BF, Skaug K, Assmus J, Bentsen SB. Effects of a COPD self-management support intervention: a randomized controlled trial. Int J Chron Obstruct Pulmon Dis 2018 Nov 8;13:3677-3688.

(200) Duggleby W, Ploeg J, McAiney C, Peacock S, Fisher K, Ghosh S, et al. Web-Based Intervention for Family Carers of Persons with Dementia and Multiple Chronic Conditions (My Tools 4 Care): Pragmatic Randomized Controlled Trial. J Med Internet Res 2018 Jun 29;20(6):e10484.

(201) Leenen LAM, Wijnen BFM, Kessels AGH, Chan H, de Kinderen RJA, Evers SMAA, et al. Effectiveness of a multicomponent self-management intervention for adults with epilepsy (ZMILE study): A randomized controlled trial. Epilepsy Behav 2018 Mar;80:259-265.

(202) ClinicalTrials.gov. DiabetesFlex - Patient Involvement and Patient-reported Outcome Measures in Type 1 Diabetes. Available at: https://clinicaltrials.gov/ct2/show/NCT03202732. Accessed 01/03, 2020.

(203) ClinicalTrials.gov. Intensified Follow-up of Lung Cancer Using Weekly Questionnaires Via the Internet (ProWide). Available at: https://clinicaltrials.gov/ct2/show/NCT03608410. Accessed 01/03, 2020.

PAPERS

Paper I:

Schougaard LMV, de Thurah A, Christensen J, Lomborg K, Maindal HT, Mejdahl CT, Vestergaard JM, Winding TN, Biering K, Hjollund NH. Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study. Qual Life Res 2020 Jan 3. (E-pub ahead of print).

Paper II:

Schougaard LMV, de Thurah A, Christiansen DH, Sidenius P, Hjollund NH. Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test-retest reliability study. BMJ Open 2018 Jul 25;8(7):e021337-2017-021337.

Paper III:

Schougaard LMV, de Thurah A, Bech P, Hjollund NH, Christiansen DH. Test-retest reliability and measurement error of the Danish WHO-5 Well-being Index in outpatients with epilepsy. Health Qual Life Outcomes 2018 Sep 6;16(1):175-018-1001-0

Paper IV:

Schougaard LMV, Mejdahl CT, Petersen KH, Jessen A, de Thurah A, Sidenius P, Lomborg K, Hjollund NH. Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: study protocol for a pragmatic randomised controlled study. BMC Health Serv Res 2017 Jan 26;17(1):83-017-2015-8

Paper V:

Schougaard LMV, Mejdahl CT, Christensen J, Lomborg K, Maindal HT, de Thurah A, Hjollund NH. Patient-initiated versus fixed-interval patient-reported outcome-based follow-up in outpatients with epilepsy: a pragmatic randomized controlled trial. J Patient Rep Outcomes 2019 Sep 13;3(1):61-019-0151-0.

PAPER I

Including online supplemental materials:

Appendix 1: Multiple imputation models

Appendix 2: Original raw analyses

Appendix 3: Sensitivity analysis

Quality of Life Research
https://doi.org/10.1007/s11136-019-02407-2

Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study

Liv Marit Valen Schougaard[1] · Annette de Thurah[2,3] · Jakob Christensen[3,4,5] · Kirsten Lomborg[3] · Helle Terkildsen Maindal[6] · Caroline Trillingsgaard Mejdahl[1] · Jesper Medom Vestergaard[7] · Trine Nøhr Winding[7] · Karin Biering[3,7] · Niels Henrik Hjollund[1,8]

Accepted: 23 December 2019

Abstract

Purpose We examined the association between sociodemographic, personal, and disease-related determinants and referral to a new model of health care that uses patient-reported outcomes (PRO) measures for remote outpatient follow-up (PRO-based follow-up).

Methods We conducted a prospective cohort study among outpatients with epilepsy at the Department of Neurology at Aarhus University Hospital, Denmark. Included were all persons aged ≥ 15 years visiting the department for the first time during the period from May 2016 to May 2018. Patients received a questionnaire containing questions about health literacy, self-efficacy, patient activation, well-being, and general health. We also collected data regarding sociodemographic status, labour market affiliation, and co-morbidity from nationwide registers. Associations were analysed as time-to-event using the pseudo-value approach. Missing data were handled using multiple imputations.

Results A total of 802 eligible patients were included in the register-based analyses and 411 patients (51%) responded to the questionnaire. The results based on data from registers indicated that patients were less likely to be referred to PRO-based follow-up if they lived alone, had low education or household income, received temporary or permanent social benefits, or if they had a psychiatric diagnosis. The results based on data from the questionnaire indicated that patients were less likely to be referred to PRO-based follow-up if they reported low levels of health literacy, self-efficacy, patient activation, well-being, or general health.

Conclusion Both self-reported and register-based analyses indicated that socioeconomically advantaged patients were referred more often to PRO-based follow-up than socioeconomically disadvantaged patients.

Keywords Patient-reported outcome measures · Ambulatory care · Outpatient clinics, hospital · Referral and consultation · Cohort study

Introduction

In 2019, it was estimated that two-thirds of the adult Danish population have one or multiple chronic conditions, a number that is expected to increase [1]. This increase contributes to a growing burden on the healthcare system, and to manage this challenge, several initiatives must be considered by the health authorities. One of these initiatives could be systematic use of patient-reported outcome (PRO) measures at the individual patient level in the healthcare system. PRO measures are defined as the patient's own report on his/her health status and symptoms without interpretation by a clinician or anyone else [2]. The use of PRO measures in individual patient management has several applications; for example, it can facilitate monitoring of symptoms before and after treatment, facilitate communication between patients and clinicians, facilitate early identification of problems, and reduce unnecessary outpatient appointments for stable patients [3, 4].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11136-019-02407-2) contains supplementary material, which is available to authorised users.

✉ Liv Marit Valen Schougaard
livschou@rm.dk

Extended author information available on the last page of the article

Published online: 03 January 2020

The Danish PRO system, AmbuFlex, is a new model for outpatient healthcare that uses PRO measures as the basis for outpatient follow-up of patients with chronic and malignant diseases [5]. The model uses PRO measures in remote outpatient follow-up in which patients report essential information about their health status and symptoms from home instead at conventional follow-up with scheduled appointments. The PRO data are used by clinicians to decide whether a patient needs or want clinical attention, making it possible to reduce the number of unnecessary outpatient appointments [6]. In addition, the model aims to improve quality of care and promote patient-centred care. In this study, remote follow-up by using PRO measures is termed PRO-based follow-up.

Since 2012, approximately 7000 outpatients with epilepsy from five Danish neurological departments have been referred to PRO-based follow-up [5, 7]. The criteria for referral to PRO-based follow-up are not defined in a standardised guideline; instead, referral is based on the individual clinicians' assessment of the patient together with the patients' preferences and capabilities. The use of PRO measures in remote outpatient follow-up is a relatively new initiative that has expanded during the last 5 years in Denmark. We have not been able to identify other PRO systems that use PRO measures as the basis for follow-up of outpatients with epilepsy or in any other outpatient population, nor studies that have investigated factors associated with patients who participate in PRO-based follow-up. However, studies regarding non-response to questionnaires have found that factors associated with non-response were lower socioeconomic status [8, 9], male sex [8, 10, 11], younger age [9, 10, 12], not living with a partner [8, 10, 13], different ethnic background than Danish [10, 12], and poorer health or quality of life [8, 9]. It is therefore important that the clinicians consider these and similar aspects when deciding whether to refer a patient to PRO-based follow-up. To the best of our knowledge, no research has been conducted to explore associations between patient characteristics and referral to PRO-based follow-up.

This study aimed to identify sociodemographic, personal, and disease-related factors associated with referral to PRO-based follow-up. We hypothesised that a low level of education and household income; higher age; solo living; passive labour market participation; a low level of health literacy, self-efficacy, patient activation, well-being, and general health; and high level of co-morbidity were associated with lower probability of referral to PRO-based follow-up.

Materials and methods

Study design, setting and participants

We conducted a prospective cohort study among outpatients with epilepsy at the Department of Neurology at Aarhus University Hospital, Denmark. All persons aged at least 15 years visiting the department for the first time between May 2016 and May 2018 with either a diagnosis or suspicion of epilepsy were invited to participate. Eligible participants were identified in the Hospital Business Intelligence (BI) Register in the Central Denmark Region, which includes information on diagnoses classified according to the international classification of disease—version 10 (ICD-10) [14]. Data were collected every second week in patients with epilepsy (DG 40–DG409), suspicion of epilepsy (DZ033A), first time unprovoked generalised seizure (DR568E), and other non-specified seizures (DR568). A physician registers all diagnoses at hospital discharge or termination of outpatient contact. The regional registers are required by law to submit standardised data to the Danish National Patient Registry (DNPR) at least monthly. The most frequently reported measure of the validity of the records in the DNPR is the positive predictive value (PPV), defined as the proportion of patients registered with a disease who truly have the disease and usually estimated using medical record review as the reference standard to confirm the presence of disease. For the diagnosis category epilepsy in the DNPR, the PPV is estimated to be 81.4% (75.2–86.3) [15].

A research questionnaire was mailed to the study participants approximately 2 weeks after their first appointment at the department. They could choose to complete either a paper- or web-based version of the questionnaire. Non-responders received one reminder after 21 days. In addition, data on all participants including responders and non-responders were obtained from regional and national registers.

Determinant variables

Data regarding cohabitation status, education, income, labour market affiliation, and co-morbidity were collected from available registers from Statistics Denmark. All Danish Citizens have a unique personal identification (CPR) number [16], which can be used to generate linkages between registers. Data regarding health literacy, self-efficacy, patient activation, well-being, and general health were collected by standardised questionnaires. Questionnaire data were linked with registry-based data from Statistics Denmark in January 2019 using patients' CPR number. Table 1 presents an overview of determinant variables and data sources.

Register data

Data on gender and age were obtained from the Hospital BI register in Central Denmark Region [14]. We used the age of the participants at the date of inclusion in the study. Age was categorised into five age groups. Cohabitation

Table 1 Overview of determinant variables and registry and questionnaire data sources

Determinant	Data source
Age Gender	The Hospital Business Intelligence (BI) Register in Central Denmark Region [14]
Cohabitation status	The Danish Civil Registration Register (CPR) [17]
Education	The Danish Education Register [18]
Household income	Danish register on income and transfer payments [19]
Labour market affiliation	The Danish Register for Evaluation and Marginalisation (DREAM) [20]
Co-morbidity Psychiatric disease	The Danish National Patient Registry (DNPR) [21]
Well-being	WHO-Five Well-Being Index (WHO-5) [28]
General health	Short Form Health Survey 36 (SF-36) [30, 31]
Health literacy	Health Literacy Questionnaire (HLQ) [23, 24]
Self-efficacy	General Self-Efficacy Scale (GSES) [25, 26]
Patient activation	Patient Activation Measure 13 (PAM-13) [32]

status was collected from The Danish Civil Registration System [17] the year before inclusion in the study and categorised into "Living with a partner/family" and "Living alone". Level of education was obtained from the Danish Education Registers [18] the year before inclusion and categorised into three groups: low (< 10 years), medium (10–12 years), or high (> 12 years) educational level. Data regarding household income were collected from the Danish registers on personal income [19] the year before inclusion and categorised into low, medium, or high income according to tertiles (33.3rd and 66.6th percentile) in the study population. If cohabitation status, education, or household income data were missing in the year before inclusion, data from the previous year were used. Information about labour market affiliation was retrieved from the Danish Register for Evaluation and Marginalisation (DREAM) [20]. DREAM is a national register which contains weekly updated information about a range of temporary and permanent social benefits. Information about labour market participation was gathered for the 52-week period before the date of inclusion in the study. Based on the amount of received benefits, the participants were divided into five groups: Self-supporting (labour market or education participation): receiving social benefits for a maximum of 4 weeks; Temporary social benefits: receiving temporary social benefits for more than 4 weeks; Permanent social benefits: receiving permanent social benefits for more than 4 weeks; and Normal retirement: receiving normal retirement benefits for more than 4 weeks. Level of co-morbidity and psychiatric diseases were extracted from the DNPR [21]. The Charlson Comorbidity Index was used to categorise the participants into three groups: 0 (Low); 1–2 (Medium); > 2 (High) level of co-morbidity [22]. Psychiatric diseases (DF 00 – 99) were dichotomised into present or not within 2 years before enrolment.

Questionnaire data

Health literacy was measured using the Health Literacy Questionnaire (HLQ), which is a multi-dimensional questionnaire measuring a broad perception of health literacy [23, 24]. The HLQ has well-documented psychometric properties [23, 24] and consists of 44 items covering nine subscales. The following subscales were used: 4: "Social support for health"; 6: "Ability to actively engage with healthcare providers"; and 9: "Understand health information well enough to know what to do". Subscale 4 has a four-point ordinal response options ranging from 1 "strongly disagree", 2 "disagree", 3 "agree" to 4 "strongly agree". Subscales 6 and 9 have a five-point ordinal response option ranging from 1 "cannot do", 2 "very difficult", 3 "quite difficult", 4 "quite easy" to 5 "very easy". The average score across all items were estimated for each of the subscales. If items were missing, the mean score of the other items were used to estimate the scale score. The score was not estimated if more than two items were missing. Higher scores indicate a higher degree of health literacy. Subscale 4 was also dichotomised to identify participants who "strongly disagree" or "disagree" (score ≤ 2) with having social support. And subscales 6 and 9 were dichotomised to identify participants who "could not" or found it "very difficult" or "quite difficult" (scores ≤ 3) to actively engage with health care providers and understand health information. Self-efficacy was measured using the General Self-Efficacy Scale (GSE), which is a 10-item questionnaire measuring optimistic self-belief to cope with difficult tasks in life [25, 26]. The psychometric properties of the scale have been evaluated in a range of different countries and populations [27]. The 10 items have four ordinal response options ranging from 1 "not at all true", 2 "hardly true", 3 "moderately true" to 4 "exactly true". The GSE score ranges from 10 to 40 (best). In addition, the GSE scale was dichotomised at the median cut-off

point in the study population: < 30 (Low) and ≥ 30 (High). WHO-Five Well-being Index (WHO-5) is a questionnaire consisting of five positively worded items reflecting current mental well-being within the previous 2 weeks [28]. The instrument has demonstrated sufficient psychometric properties in a wide range of chronic conditions [28, 29]. Items are rated on a six-point ordinal scale ranging from 5 "all of the time", 4 "most of the time", 3 "more than half of the time", 2 "less than half of the time", 1 "some of the time" to 0 "at no time". The score ranges from 0 to 100. Higher scores indicate a better degree of well-being, and a score below 50 indicates increased risk of depression [28]. The WHO-5 score was also dichotomised at < 50 (low) and ≥ 50 (high). The GSE and WHO-5 scores were not estimated if there were missing items. The Short Form Health Survey (SF-36) is a multi-dimensional questionnaire with eight subscales measuring different aspects of physical and mental health [30, 31]. In this study, only one single item was included: "In general, would you say your health is: excellent, very good, good, fair, or poor". The variable was divided into three groups: "excellent/very good", "good", and "fair/poor". Patient Activation Measure 13 (PAM-13) is a 13-item questionnaire measuring the aspect patient activation in health [32]. In this study, only two single items were included, which were modified from the PAM scale: "I am confident that I can tell when I need to get outpatient care" and "I am confident I can figure out solutions when new situations or problems arise with my health condition", with the response categories: "disagree strongly", "disagree", "agree", and "agree strongly". The item responses were dichotomised into "disagree strongly/disagree" and "agree/agree strongly".

Outpatient follow-up

Department of Neurology at Aarhus University Hospital offers both PRO-based and conventional follow-up. In PRO-based follow-up, outpatients receive fixed-interval disease-specific questionnaires instead of in-clinic visits [5]. The questionnaire is coupled with a pre-defined color-algorithm used to determine whether the patients need clinical attention. Green color indicates no need of attention, red color indicates need of attention, and yellow color indicates that the patient might need attention [5, 7]. Clinicians assess the questionnaire responses together with other relevant data from the Electronic Health Record [5]. As of December 2019, 2110 epilepsy outpatients (approximately 50% of the entire outpatient population) are attending PRO-based follow-up at the department. In conventional follow-up, patients receive in-clinic visits or telephone consultations. A clinician together with the patient decides in each case whether to refer to PRO-based follow-up or conventional follow-up.

Outcome

Patients were followed up from their first visit at the Department of Neurology at Aarhus University Hospital (the date of inclusion) until the event of interest: referral to PRO-based follow-up. Patients were censored at study end (after 18 months' follow-up or in January 2019) or in the event of no further need for outpatient follow-up, emigration, or death, whichever came first. Patients referred to PRO-based follow-up are registered in the AmbuFlex-system by a clinician [5]. We used the date of this registration to define whether a patient was referred to PRO-based follow-up. The date was gathered from the AmbuFlex-database [5, 6]. The dates of other events (no further need for outpatient follow-up, emigration, or death) were obtained from the Hospital BI Register in the Central Denmark Region.

Statistical analysis

We analysed the associations between register- and questionnaire-based determinants and the proportion of patients referred to PRO-based follow-up within 6, 12, and 18 months after the patients' first visit at the department. Not all participants were followed for 12 and 18 months; therefore, the analyses were based on time to event by using the pseudo-value approach to examine the cumulative risk ratio (RR) at the three time points [33, 34]. In the pseudo-value approach, pseudo values are generated and used in a generalised linear regression. Death, emigration, and end of follow-up were considered competing risks in the model if they occurred before the event of interest. All estimates were reported with 95% confidence intervals. Age, gender, cohabitation status, education, and co-morbidity were included in the adjusted analyses. The confounder variables were selected a priori based on associations between these factors and questionnaire non-response in previous studies [8–13].

To manage the missing data problem, we decided to use the multiple imputation method [35]. Based on the assumption that data were missing at random, 100 complete datasets were created based on a model of all relevant variables measured in the population (Appendix 1). The robustness of the imputed model was evaluated by modifying the variables in the model. Furthermore, sensitivity analyses were performed in which we assumed that data were not missing at random; for example, if data were missing in patients with lower health literacy than expected from the imputations. For this group of patients, the imputed health literacy scores were reduced with one point corresponding to approximately one standard deviation (SD). Thereafter, the cumulative RRs at the three time points were analysed in a generalised linear regression using the pseudo-value approach.

Categorical data were presented as numbers and percentages. For normally distributed continuous data, means and

SDs were presented, and for non-normally distributed questionnaire data, median and interquartile ranges were also presented. Data collected from registers were used to compare non-responders of questionnaire data with responders. All analyses were performed using STATA version 15 (Stata Corporation, College Station, Texas, USA).

Results

Characteristics of the study population

From May 2016 to May 2018, a total of 822 patients had their first visit at the Department of Neurology at Aarhus University Hospital with either a diagnosis or suspicion of epilepsy (Fig. 1). Twenty patients were excluded due to termination of outpatient care, emigration, or death before start of follow-up, leaving 802 patients in the study. The mean age of the study population was 49.3 years (SD 21.9 years) and 52% were male (Table 2). Only 13% had a high level of co-morbidity (Charlson Comorbidity Index > 2) and 12% had a psychiatric disease diagnosis. Data were missing for three register-based variables: cohabitation status (2%), education level (7%), and household income (1%). The overall response rate was 51%, 61% for patients referred to PRO-based follow-up and 48% for patients not referred ($p = 0.003$). Questionnaire non-responders were younger ($p < 0.001$), lower educated ($p = 0.03$), more likely lived alone ($p < 0.001$), had lower household income ($p = 0.01$), and received more often temporary or social benefits ($p < 0.001$) than responders. No differences were found with regard to gender, co-morbidity, and psychiatric disease. Table 3 presents an overview of the self-reported questionnaire data from the 411 responders. There were fewer than 5% missing items for all scales expect for the GSE scale, where 9% of items were missing. The stratified data according to 18-month follow-up status

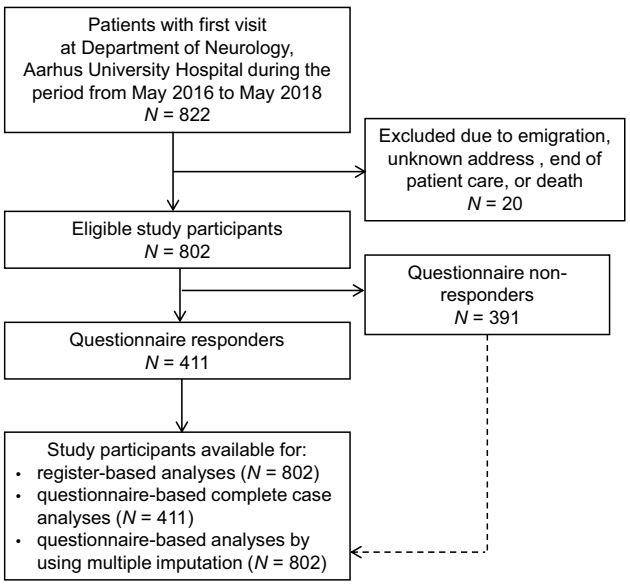


Fig. 1 Flowchart of patients included in the study

Table 2 Baseline register characteristics of 802 patients and among questionnaire responders and non-responders from the Department of Neurology, Aarhus University Hospital May 2016–May 2018

	Total *n* (%)	Responders *n* (%)	Non-responders *n* (%)
	N = 802	*N* = 411	*N* = 391
Age, years			
15–24	168 (21)	78 (19)	90 (23)
25–39	136 (17)	46 (11)	90 (23)
40–59	187 (23)	93 (23)	94 (24)
60–69	124 (15)	76 (18)	48 (12)
70–99	187 (23)	118 (29)	69 (18)
Gender			
Female	387 (48)	204 (50)	183 (47)
Male	415 (52)	207 (50)	208 (53)
Cohabitation status			
Not living alone	492 (61)	280 (68)	212 (54)
Solo living	290 (36)	124 (30)	166 (42)
Missing	20 (2)	7 (2)	13 (3)
Education			
High (> 12 years)	135 (17)	79 (19)	56 (14)
Medium (10–12 years)	281 (35)	154 (37)	127 (32)
Low (< 10 years)	327 (41)	154 (37)	173 (44)
Missing	59 (7)	24 (6)	35 (9)
Household income[a]			
High	266 (33)		
Medium	267 (33)		
Low	261 (33)		
Missing	8 (1)		
Labour market affiliation			
Self-supporting	243 (30)	132 (32)	111 (28)
Normal retirement	234 (29)	151 (37)	83 (21)
Temporary social benefits	174 (22)	69 (17)	105 (27)
Permanent social benefits	151 (19)	59 (14)	92 (24)
Co-morbidity (Charlson index)			
Low 0	463 (58)	231 (56)	232 (59)
Medium 1–2	233 (29)	122 (30)	111 (28)
High > 2	106 (13)	58 (14)	48 (12)
Psychiatric disease			
No	703 (88)	365 (89)	338 (86)
Yes	99 (12)	46 (11)	53 (14)

[a]According to guidelines from Statistics Denmark, the distribution of household income was not reported for responders and non-responders because the missing number is below five observations

indicated that patients who received conventional follow-up reported lower levels of all measured constructs than did patients referred to PRO-based follow-up.

Mean follow-up time was 10.6 months (SD 6.6 months). At 6-, 12-, and 18-month follow-up, 139, 173, and 185 patients had been referred to PRO-based follow-up, 95, 129, and 172 patients were no longer visiting the outpatient clinic, and 26, 43, and 52 patients had died, respectively. Fewer than 5 patients had emigrated at all three time points.

Register determinants of referral to PRO-based follow-up

The cumulative risk ratios of referral to PRO-based follow-up 6, 12, and 18 months after the patients' first visit at the department in relation to register determinants are presented in Table 4. At all three time points, a decreased adjusted risk of referral to PRO-based follow-up was found among patients who lived alone and patients with low education or household income. At 18-month follow-up, a decreased adjusted risk was also found in patients with temporary or permanent social benefits and patients with a psychiatric diagnosis. Further, at 18-month follow-up, men were more likely to be referred to PRO-based follow-up than women. At 12-month follow-up, a decreased adjusted risk of referral was found in patients with a medium level of co-morbidity compared to a low level of co-morbidity; however, no differences were found at 6- or 18-month follow-up. No differences were found between age groups.

Questionnaire determinants of referral to PRO-based follow-up

The cumulative risk ratios of referral to PRO-based follow-up 6, 12, and 18 months after the patients' first visit at the department in relation to questionnaire determinants are presented in Tables 5 and 6. Patients who reported a low level of perceived confidence regarding to figure out solutions or problems related to their health condition had a decreased adjusted risk of referral to PRO-based follow-up at all three time points. At 12- and 18-month follow-up, a decreased adjusted risk was found in patients with low health literacy (HLQ 9), well-being, or general health, and in patients who reported a low level of perceived confidence to decide their need for outpatient care. At 18-month follow-up, a decreased adjusted risk of referral was also found in patients with low self-efficacy. No adjusted differences were found in health literacy (HLQ 4 and HLQ 6). The questionnaire scale scores were also analysed (Table 6). In the adjusted analyses, we found that a one-unit increase in mean scale scores of HLQ 6 and 9 increased the risk of referral to PRO-based follow-up at all three time points. Similarly, at 12- and 18-month follow-up, one-unit increase in mean scale scores of HLQ 4 and WHO-5 also increased the risk of referral to PRO-based follow-up. One-unit increase in mean scale score of GSE increased the risk of referral at 18-month follow-up.

Other analyses

The results based on multiple imputations were comparable to the original raw data analyses (Appendix 2). The imputations did not change the estimates markedly. In addition, the results were not affected by modifying the variables in the multiple imputation model. The sensitivity analyses of the health literacy scores did not alter the results noticeably (Appendix 3).

Discussion

This study showed that several sociodemographic, personal, and disease-related factors play a role in referral to PRO-based follow-up. Patients were less likely to be referred to PRO-based follow-up if they lived alone or had low education or household income, if they received temporary or permanent social benefits, or if they had a psychiatric diagnosis. Furthermore, we found that patients were less likely to be referred to PRO-based follow-up if they reported a low level of health literacy, self-efficacy, patient activation, well-being, or general health.

A shift toward more active involvement of patients in chronic disease care management is taking place in the healthcare system, characterised by productive interactions between patients and health care providers [36, 37]. These interactions do not necessarily require face-to-face visits [37]. In addition, management of chronic diseases is shifting from the clinic to the patients' homes [38]. PRO-based follow-up and telephone consultations provided by nurses [39] are examples of care at a distance in epilepsy outpatient follow-up. John et al. argue that telephone follow-up could replace traditional scheduled appointments unless there is a clear clinical need; for example, if the patient is considered vulnerable [39].

Patients with chronic diseases need the skills, confidence, and information necessary to make best use of their involvement in self-management [37]. They must cope with increasing treatment workload, e.g. taking medication, reading information, and making lifestyle changes. The treatment workload should balance the patients' capacity [38, 40]. Low capacity concurrent with high workload may diminish self-management, adherence to treatment, and health outcomes [38]. Increased risk of depression, poorer quality of life, and more social stigma have been found in patients with seizures compared to patients with no seizures [41, 42], which supports the need for tighter follow-up strategy in patients with severe epilepsy than in patients with stable

Table 3 Baseline self-reported characteristics and stratified according to status at 18 month follow-up[a] among 411 patients from the Department of Neurology, Aarhus University Hospital May 2016–May 2018

	Total (%)	PRO-based follow-up	Conventional follow-up	End of outpatient care or emigrated
	N = 411	*n* = 113	*n* = 197	*n* = 82
Social support for health (HLQ4)				
Mean (SD)	3.3 (0.60)	3.4 (0.49)	3.2 (0.63)	3.2 (0.67)
Median (IQR)	3.4 (3.0−3.8)	3.4 (3.0−3.8)	3.2 (3.0−3.6)	3.4 (3.0−3.8)
Missing, *n* (%)	10 (2.4)			
Ability to actively engage with healthcare providers (HLQ6)				
Mean (SD)	3.6 (0.95)	3.9 (0.83)	3.5 (0.98)	3.6 (0.99)
Median (IQR)	3.8 (3.0−4.2)	4.0 (3.4−4.4)	3.6 (3.0−4.2)	3.8 (3.1−4.3)
Missing, *n* (%)	9 (2.2)			
Understanding health information well enough to know what to do (HLQ9)				
Mean (SD)	3.6 (0.97)	3.9 (0.78)	3.5 (1.02)	3.7 (0.95)
Median (IQR)	3.8 (3.0−4.3)	4.0 (3.5−4.4)	3.8 (2.8−4.2)	4.0 (3.2−4.4)
Missing, *n* (%)	9 (2.2)			
Self-efficacy (GSE)				
Mean (SD)	27.4 (7.4)	29.0 (5.9)	26.0 (7.8)	29.5 (6.6)
Median (IQR)	29 (23−33)	30 (26−32)	27 (20−32)	30 (26−34)
Missing, *n* (%)	37 (9.0)			
Well-being (WHO-5)				
Mean (SD)	61.3 (23.9)	66.8 (21.0)	56.2 (24.9)	66.9 (20.4)
Median (IQR)	64 (48−80)	72 (56−80)	60 (36−76)	72 (52−80)
Missing, n (%)	17 (4.1)			
General health, *n* (%)				
Excellent	36 (8.8)			
Very good	94 (22.9)			
Good	149 (36.3)			
Fair	86 (20.9)			
Poor	33 (8.0)			
Missing	13 (3.1)			
Patient activation[b], *n* (%)				
Disagree strongly	31 (7.5)			
Disagree	102 (24.8)			
Agree	170 (41.4)			
Agree strongly	94 (22.9)			
Missing	14 (3.4)			
Patient activation[c], *n* (%)				
Disagree Strongly	49 (11.9)			
Disagree	86 (20.9)			
Agree	190 (46.2)			
Agree strongly	74 (18.0)			
Missing	12 (2.9)			

SD standard deviation; *IQR* interquartile range; *HLQ* Health Literacy Questionnaire; *GSE* General Self-efficacy scale; *WHO-5* WHO-Five Well-being Index

[a]According to Statistics Denmark's guidelines, the distribution of variables was not reported if a cell contained less than five observations. Among 411 patients, 19 patients had died at 18-month follow (data not shown)

[b]I am confident that I can tell when I need to get outpatient care

[c]I am confident I can figure out solutions when new situations or problems arise with my health condition

Table 4 Risk ratio (RR) of referral to PRO-based follow-up 6, 12, and 18 months after the first visit at Department of Neurology, Aarhus University Hospital according to register determinants ($N = 802$)

	6-month follow-up		12-month follow-up		18-month follow-up	
	Crude RR	Adjusted RR[a]	Crude RR	Adjusted RR[a]	Crude RR	Adjusted RR[a]
Age, years						
15–24	Ref	Ref	Ref	Ref	Ref	Ref
25–39	1.04 (0.63–1.70)	0.99 (0.57–1.71)	0.91 (0.59–1.40)	0.85 (0.54–1.36)	0.90 (0.59–1.36)	0.84 (0.54–1.30)
40–59	1.08 (0.68–1.71)	0.98 (0.58–1.65)	1.02 (0.69–1.50)	0.94 (0.60–1.46)	1.01 (0.70–1.47)	0.92 (0.60–1.40)
60–69	1.08 (0.65–1.80)	0.77 (0.43–1.37)	0.98 (0.63–1.52)	0.83 (0.50–1.37)	0.85 (0.55–1.30)	0.73 (0.45–1.19)
70–99	1.01 (0.64–1.60)	0.92 (0.52–1.62)	0.90 (0.60–1.34)	1.06 (0.62–1.78)	0.84 (0.57–1.24)	0.95 (0.58–1.58)
Gender						
Female	Ref	Ref	Ref	Ref	Ref	Ref
Male	1.32 (0.97–1.80)	1.35 (0.95–1.92)	1.29 (0.99–1.69)	1.28 (0.96–1.69)	1.39 (1.07–1.81)	1.36 (1.03–1.79)
Cohabitation status						
Living with a partner/family	Ref	Ref	Ref	Ref	Ref	Ref
Living alone	0.61 (0.43–0.86)	0.60 (0.41–0.89)	0.55 (0.40–0.75)	0.55 (0.39–0.78)	0.58 (0.43–0.79)	0.63 (0.45–0.89)
Education						
High (> 12 years)	Ref	Ref	Ref	Ref	Ref	Ref
Medium (10–12 years)	1.03 (0.70–1.51)	0.97 (0.66–1.46)	1.25 (0.88–1.78)	1.22 (0.86–1.74)	1.20 (0.85–1.69)	1.15 (0.82–1.62)
Low (< 10 years)	0.53 (0.34–0.81)	0.46 (0.28–0.75)	0.65 (0.44–0.96)	0.62 (0.41–0.94)	0.69 (0.47–1.01)	0.65 (0.44–0.97)
Household income						
High	Ref	Ref	Ref	Ref	Ref	Ref
Medium	0.46 (0.32–0.67)	0.61 (0.39–0.94)	0.52 (0.38–0.71)	0.69 (0.48–1.00)	0.50 (0.37–0.68)	0.64 (0.44–0.93)
Low	0.49 (0.34–0.71)	0.59 (0.38–0.90)	0.43 (0.30–0.60)	0.51 (0.35–0.75)	0.41 (0.29–0.57)	0.47 (0.32–0.68)
Labour market affiliation						
Self-supporting	Ref	Ref	Ref	Ref	Ref	Ref
Normal retirement	0.71 (0.49–1.01)	0.69 (0.42–1.13)	0.74 (0.54–1.02)	0.86 (0.55–1.37)	0.65 (0.48–0.89)	0.77 (0.49–1.20)
Temporary social benefits	0.69 (0.47–1.03)	0.68 (0.45–1.05)	0.78 (0.55–1.09)	0.79 (0.56–1.13)	0.68 (0.49–0.95)	0.69 (0.49–0.96)
Permanent social benefits	0.38 (0.22–0.67)	0.55 (0.29–1.04)	0.44 (0.27–0.71)	0.60 (0.35–1.02)	0.39 (0.24–0.62)	0.51 (0.31–0.84)
Co-morbidity (Charlson Index)						
Low 0	Ref	Ref	Ref	Ref	Ref	Ref
Medium 1–2	0.79 (0.54–1.14)	0.73 (0.48–1.12)	0.72 (0.52–1.00)	0.69 (0.46–0.97)	0.74 (0.54–1.01)	0.74 (0.52–1.04)
High > 2	1.09 (0.72–1.66)	1.08 (0.68–1.71)	0.88 (0.60–1.29)	0.89 (0.60–1.33)	0.78 (0.53–1.14)	0.82 (0.55–1.23)
Psychiatric disease						
No	Ref	Ref	Ref	Ref	Ref	Ref
Yes	0.65 (0.37–1.15)	0.75 (0.42–1.36)	0.47 (0.26–0.85)	0.55 (0.30–1.01)	0.44 (0.24–0.82)	0.50 (0.27–0.93)

Numbers in round brackets are 95% confidence intervals (CIs). The estimated RRs and 95% CIs were obtained after multiple imputations in a generalised linear regression using the pseudo-value approach

[a]Mutual adjusted for age, gender, cohabitation status, education, and co-morbidity

disease. PRO-based follow-up aims to optimise the healthcare resources, as patients with no need of clinical attention are not routinely seen in-clinic. Hence, resources can be used to respond rapidly to patients with a high symptom burden. A qualitative study found that clinicians experienced that problems were more complex in the patients seen in-clinic after implementation of PRO-based follow-up [43].

We found that PRO-based follow-up is offered to a selected group of socioeconomically advantaged patients. The goal has never been to refer the whole outpatient population. The decision must be based on the patient's preferences and clinical profile. In PRO-based follow-up, patients fill in scheduled questionnaires during follow-up; hence, we considered it relevant to consider factors related to questionnaire non-response. We found that lower sociodemographic status was associated with a decreased probability of referral to PRO-based follow-up. This finding is supported by studies regarding questionnaire non-response [8, 9]. A Danish study among patients with endometrial cancer found that well-educated patients more often sought medical attendance if symptoms of recurrence occurred than did less educated patients [44]. We also found an association between a lower

Table 5 Risk ratio (RR) of referral to PRO-based follow-up 6, 12, and 18 months after the first visit at Department of Neurology, Aarhus University Hospital according to questionnaire determinants ($N = 802$)

	6-month follow-up		12-month follow-up		18-month follow-up	
	Crude RR	Adjusted RR[a]	Crude RR	Adjusted RR[a]	Crude RR	Adjusted RR[a]
Social support for health (HLQ4)						
High (> 2)	Ref	Ref	Ref	Ref	Ref	Ref
Low (≤ 2)	0.47 (0.09–2.44)	0.43 (0.07–2.48)	0.43 (0.09–2.02)	0.40 (0.08–2.06)	0.38 (0.07–2.02)	0.40 (0.07–2.20)
Ability to actively engage with healthcare providers (HLQ6)						
High (> 3)	Ref	Ref	Ref	Ref	Ref	Ref
Low (≤ 3)	0.60 (0.37–0.98)	0.73 (0.42–1.28)	0.55 (0.35–0.88)	0.63 (0.38–1.05)	0.55 (0.35–0.89)	0.64 (0.39–1.05)
Understanding health information well enough to know what to do (HLQ9)						
High (> 3)	Ref	Ref	Ref	Ref	Ref	Ref
Low (≤ 3)	0.42 (0.24–0.75)	0.53 (0.28–1.01)	0.41 (0.23–0.71)	0.48 (0.26–0.87)	0.40 (0.23–0.69)	0.45 (0.25–0.82)
Self-efficacy (GSE)						
High (≥ 30)	Ref	Ref	Ref	Ref	Ref	Ref
Low (< 30)	0.61 (0.42–0.89)	0.69 (0.46–1.04)	0.63 (0.46–0.88)	0.73 (0.52–1.03)	0.60 (0.43–0.84)	0.69 (0.49–0.98)
Well-being (WHO-5)						
High (≥ 50)	Ref	Ref	Ref	Ref	Ref	Ref
Low (< 50)	0.67 (0.43–1.03)	0.73 (0.45–1.18)	0.54 (0.36–0.82)	0.59 (0.38–0.91)	0.50 (0.33–0.78)	0.55 (0.35–0.86)
General health						
Excellent/very good	Ref	Ref	Ref	Ref	Ref	Ref
Good	1.02 (0.67–1.54)	1.07 (0.69–1.65)	0.82 (0.56–1.19)	0.88 (0.60–1.30)	0.71 (0.50–1.02)	0.76 (0.52–1.12)
Fair/poor	0.61 (0.36–1.02)	0.72 (0.41–1.25)	0.47 (0.29–0.75)	0.57 (0.34–0.94)	0.38 (0.24–0.61)	0.46 (0.28–0.76)
Patient activation[b]						
Agree strongly/agree	Ref	Ref	Ref	Ref	Ref	Ref
Disagree strongly/disagree	0.55 (0.36–0.86)	0.64 (0.38–1.07)	0.57 (0.38–0.85)	0.65 (0.42–0.99)	0.51 (0.34–0.76)	0.57 (0.38–0.88)
Patient activation[c]						
Agree strongly/agree	Ref	Ref	Ref	Ref	Ref	Ref
Disagree strongly/disagree	0.54 (0.34–0.86)	0.58 (0.34–0.97)	0.56 (0.37–0.86)	0.59 (0.37–0.93)	0.50 (0.33–0.76)	0.54 (0.35–0.83)

HLQ Health Literacy Questionnaire; *GSE* General Self-efficacy scale; *WHO-5* WHO-Five Well-being Index

Numbers in round brackets are 95% confidence intervals (CIs).The estimated RRs and 95% CIs were obtained after multiple imputations in a generalised linear regression using the pseudo-value approach

[a]Adjusted for age, gender, cohabitation status, education, and co-morbidity

[b]I am confident that I can tell when I need to get outpatient care

[c]I am confident I can figure out solutions when new situations or problems arise with my health condition

degree of self-reported patient activation and non-referral to PRO-based follow-up. A recent study of patients referred to advanced heart failure therapy used the PAM scale to measure the degree of patient activation [45]. In accordance with our findings, they also found that those not selected for therapy were more likely to have lower patient activation than those who were selected [45].

A qualitative study has documented a variation in patients' preferences for being active and taking responsibility in PRO-based follow-up, as some patients experienced a lack of confidence in their own capability to participate [46]. In addition, a study regarding the clinician perspective indicated that some clinicians had concerns regarding some patients' capability to participate in PRO-based follow-up, even though the patient had already been referred [43]. We cannot rule out that some clinicians were more reluctant to introduce PRO-based follow-up and did not refer all relevant patients. On the other hand, patients could also have been referred without having the skills or confidence to participate. Preferably, referral to PRO-based follow-up should be based on a shared decision between the patient and the clinician in which both advantages and disadvantages are discussed. This may strengthen the patients' expectations and willingness to participate in PRO-based follow-up and prevent non-response and dropout during PRO-based follow-up.

A high level of health literacy skills has been associated with health-promoting behaviours and better health outcomes in relation to self-reported health status, dietary habits, physical activity, smoking, alcohol consumption, and glycaemic control of diabetes [47–49]. We found that lower

Table 6 Risk ratio (RR) of referral to PRO-based follow-up 6, 12, and 18 months after the first visit at Department of Neurology, Aarhus University Hospital according to questionnaire scale scores ($N = 802$)

	6-month follow-up		12-month follow-up		18-month follow-up	
	Crude RR	Adjusted RR[a]	Crude RR	Adjusted RR[a]	Crude RR	Adjusted RR[a]
Social support for health (HLQ4 score)	1.40 (1.04–1.87)	1.35 (0.99–1.87)	1.38 (1.06–1.80)	1.31 (1.00–1.72)	1.49 (1.13–1.97)	1.40 (1.05–1.87)
Ability to actively engage with healthcare providers (HLQ6 score)	1.35 (1.12–1.64)	1.26 (1.00–1.59)	1.38 (1.16–1.65)	1.32 (1.09–1.61)	1.40 (1.18–1.66)	1.34 (1.12–1.61)
Understanding health information well enough to know what to do (HLQ9 score)	1.39 (1.15–1.67)	1.30 (1.04–1.63)	1.35 (1.16–1.57)	1.30 (1.09–1.54)	1.40 (1.19–1.65)	1.36 (1.13–1.62)
Self-efficacy (GSE score)	1.03 (1.01–1.05)	1.02 (0.99–1.05)	1.03 (1.01–1.05)	1.02 (0.99–1.04)	1.03 (1.01–1.06)	1.03 (1.00–1.05)
Well-being Index (WHO-5 score)	1.01 (1.00–1.02)	1.01 (0.99–1.01)	1.01 (1.00–1.02)	1.02 (1.00–1.02)	1.01 (1.01–1.02)	1.01 (1.00–1.02)

HLQ Health Literacy Questionnaire; *GSE* General Self-efficacy scale; *WHO-5* WHO-Five Well-being Index

Numbers in round brackets are 95% confidence intervals (CIs).The estimated RRs and 95% CIs were obtained after multiple imputations in a generalised linear regression using the pseudo-value approach

[a]Adjusted for age, gender, cohabitation status, education, and co-morbidity

health literacy was associated with a decreased probability of referral to PRO-based follow-up. Although our finding indicated that clinicians are aware of the patients' health literacy before referring a patient to PRO-based follow-up, a future focus should be on how healthcare services can be supportive toward vulnerable patients. Follow-up care for patients with a low level of health literacy may need to be more clinician-driven, clinicians needing to concentrate on increasing the health literacy level to prevent diminished participation in activities in relation to disease prevention or progression. PRO measures during follow-up of vulnerable patients may be difficult because some patients with epilepsy have cognitive disabilities and are not capable of filling in a questionnaire on their own. For this reason, the department has developed a PRO-based proxy solution in which a relative or social worker fills in the questionnaire on behalf of the patient. As of December 2019, 85 patients are attending the proxy solution at the department.

This is a large Danish prospective cohort study among outpatients with epilepsy in which register data from all the included participants were used. The study population was identified in the Hospital BI Register in the Central Denmark Region by using four selected ICD-10 codes. The selection of codes was based on a random sample of epilepsy outpatients attending PRO-based follow-up in 2015 and advice from a neurologist at the department. The four codes covered 96.4% of the diagnoses given to the patients in the random sample. Although the PPV of epilepsy diagnoses was high in the DNPR, the completeness of the four ICD-10 codes in the register is unknown; hence, there may have been patients with epilepsy or suspicion of epilepsy at the department who were not recorded in the database. However, lack of registration in the DNPR or the BI register is considered to be random; thus, the risk of bias related to selection of the study population is considered to be limited. Information bias related to registry-based analyses is also considered to be limited. Misclassification of data from registers is most likely random since the data collection is based on administrative requirements. Any potential bias would be non-differential as any missing data or misclassification took place before the event of interest (PRO-based follow-up).

Questionnaire non-response could potentially bias the estimates in both directions. The response rate was only 51% and non-responders differed from responders, as they were younger, lower educated, and received more temporary social benefits. Questionnaire non-responders were also related to the event of interest as they were less likely to be referred to PRO-based follow-up. We assumed that data were missing at random, but it is not possible to prove this. Because data may not have been missing at random, we assumed in the sensitivity analyses that the scores of HLQ were lower than the imputed values for patients with missing HLQ scores. However, the results did not change noticeably. The risk of information bias should also be considered for self-reported data. The questionnaire response took place before referral to PRO-based follow-up; thus, any misclassification of self-reported information most likely resulted in non-differential bias. We decided to dichotomise the questionnaire scale scores into a low or high level of the construct of interest to better interpret and present the results. However, dichotomisation of continuous variables entails loss of information and statistical power [50]. As can be seen in the wide confidence intervals of HLQ4 in Table 5, few participants reported 'disagree or disagree strongly' to the questions regarding social support for health. Thus, in this

study, the continuous HLQ4 scale contained more information and statistical power than the dichotomised form.

The study population was recruited from only one neurologic department in Denmark. The department is a large, highly specialised department with a large number of epilepsy patients compared to minor regional hospitals. The department was also the first department in Denmark to offer PRO-based follow-up for outpatients with epilepsy, and has a long experience with the use of PRO measures in remote outpatient follow-up. However, despite this being a single-unit study, we expect that the results may be generalised to outpatients with epilepsy and perhaps also to other patient populations with a chronic or long-term condition.

Conclusion

PRO-based follow-up has been used in Denmark since 2012, and since then, approximately 7000 epilepsy outpatients have been referred to PRO-based follow-up at five hospitals. Several sociodemographic, personal, and disease-related factors play a role in referral to PRO-based follow-up. Both register and questionnaire data were consistent and indicated that socioeconomically advantaged patients were more likely to be referred to PRO-based follow-up than less socioeconomically advantaged patients. Further research should explore how health care services to a larger extent can be supportive towards less advantaged patients.

Authors' contribution Study design: LMVS, AdT, HTM, and NHH. Data collection, management, and analyses: LMVS, NHH, JMV, JC, TNW, and KB. All authors were involved in the manuscript preparation and all authors read and approved the final manuscript.

Funding The study was funded by Aarhus University, Central Denmark Regions Health Research Foundation, and TrygFonden. JC received research support from the Danish Epilepsy Association, Central Denmark Region, and the Novo Nordisk Foundation (NNF16OC0019126).

Compliance with ethical standards

Conflict of interest JC received honoraria for serving on the scientific advisory board of UCB Nordic and Eisai AB and for giving lectures for UCB Nordic and Eisai AB. JC also received funding for a trip from UCB Nordic. The other authors have no competing interests.

Ethical approval All procedures performed in this study were in accordance with Danish ethical standards and with the Helsinki Declaration. The Danish Data Protection Agency has approved the study (reference number 1-16-02-691-14). The Ethics Committee of Central Denmark Region was contacted. The committee subsequently stated that the present study does not require approval. According to Danish law, questionnaire and register-based studies only need approval from the committee if the data include human biological material (§ 14) [51]. The correspondence with the ethical committee can be obtained from the authors on request. All data were stored and treated with confidentiality. Results from Statistic Denmark are only accessible for the researcher at an aggregated level, not at the individual patient level.

Informed consent Consent to participate in the study was informed, specific, voluntary, and explicit in accordance with guidelines from the Danish Data Protection Agency [52]. The information to the patient was provided in a letter together with the questionnaire. The patients were informed that if they responded to the questionnaire, they gave active consent for their participation and use of their data in the study. The patient was furthermore informed that the consent of participation could be withdrawn at any time.

References

1. Hvidberg, M. F., Johnsen, S. P., Davidsen, M., & Ehlers, L. (2019). A nationwide study of prevalence rates and characteristics of 199 chronic conditions in Denmark. *PharmacoEconomics-open.*
2. US Department of Health and Human Services Food and Drug Administration. (2009). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. Retrieved February 20 2019 from www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.
3. Calvert, M., Kyte, D., Price, G., Valderas, J. M., & Hjollund, N. H. (2019). Maximising the impact of patient reported outcome assessment for patients and society. *BMJ (Clinical Research ed.), 364,* k5267.
4. Greenhalgh, J. (2009). The applications of PROs in clinical practice: what are they, do they work, and why? *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 18*(1), 115–123.
5. Schougaard, L. M., Larsen, L. P., Jessen, A., Sidenius, P., Dorflinger, L., de Thurah, A., et al. (2016). AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 25*(3), 525–534.
6. Hjollund, N. H., Larsen, L. P., Biering, K., Johnsen, S. P., Riiskjaer, E., & Schougaard, L. M. (2014). Use of patient-reported outcome (PRO) measures at group and patient levels: Experiences from the generic integrated PRO system, WestChronic. *Interactive Journal of Medical Research, 3*(1), e5.
7. Schougaard, L. M. V., de Thurah, A., Christiansen, D. H., Sidenius, P., & Hjollund, N. H. (2018). Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: A test-retest reliability study. *BMJ Open, 8*(7), e021337-2017-021337.
8. Bronserud, M. M., Iachina, M., Green, A., Groenvold, M., Dorflinger, L., & Jakobsen, E. (2019). Patient-reported outcomes

(PROs) in lung cancer: Experiences from a nationwide feasibility study. *Lung Cancer (Amsterdam, Netherlands), 128,* 67–73.
9. Multone, E., Vader, J. P., Mottet, C., Schoepfer, A., Fournier, N., Burnand, B., et al. (2015). Characteristics of non-responders to self-reported questionnaires in a large inflammatory bowel disease cohort study. *Scandinavian Journal of Gastroenterology, 50*(11), 1348–1356.
10. Christensen, A. I., Ekholm, O., Gray, L., Glumer, C., & Juel, K. (2015). What is wrong with non-respondents? Alcohol-, drug- and smoking-related mortality and morbidity in a 12-year follow-up study of respondents and non-respondents in the Danish health and morbidity survey. *Addiction (Abingdon, England), 110*(9), 1505–1512.
11. Abrahamsen, R., Svendsen, M. V., Henneberger, P. K., Gundersen, G. F., Toren, K., Kongerud, J., et al. (2016). Non-response in a cross-sectional study of respiratory health in Norway. *BMJ open, 6*(1), e009912-2015-009912.
12. Kjaer, A. S. H. K., Rasmussen, T. A., Hjollund, N. H., Rodkjaer, L. O., & Storgaard, M. (2018). Patient-reported outcomes in daily clinical practise in HIV outpatient care. *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases, 69,* 108–114.
13. Jamilloux, Y., Sarabi, M., Kerever, S., Boussely, N., le Sidaner, A., Valgueblasse, V., et al. (2015). Adherence to online monitoring of patient-reported outcomes by patients with chronic inflammatory diseases: A feasibility study. *Lupus, 24*(13), 1429–1436.
14. Business Intelligence in Central Denmark Region [In Danish]. Retrieved August 27 2019 from https://www.rm.dk/om-os/organisation/it/business-intelligence/.
15. Schmidt, M., Schmidt, S. A., Sandegaard, J. L., Ehrenstein, V., Pedersen, L., & Sorensen, H. T. (2015). The Danish national patient registry: A review of content, data quality, and research potential. *Clinical Epidemiology, 7,* 449–490.
16. Frank, L. (2000). Epidemiology: When an entire country is a cohort. *Science (New York, N.Y.), 287*(5462), 2398–2399.
17. Pedersen, C. B. (2011). The Danish civil registration system. *Scandinavian Journal of Public Health, 39*(7 Suppl), 22–25.
18. Jensen, V. M., & Rasmussen, A. W. (2011). Danish education registers. *Scandinavian Journal of Public Health, 39*(7 Suppl), 91–94.
19. Baadsgaard, M., & Quitzau, J. (2011). Danish registers on personal income and transfer payments. *Scandinavian Journal of Public Health, 39*(7 Suppl), 103–105.
20. Hjollund, N. H., Larsen, F. B., & Andersen, J. H. (2007). Register-based follow-up of social benefits and other transfer payments: accuracy and degree of completeness in a Danish interdepartmental administrative database compared with a population-based survey. *Scandinavian Journal of Public Health, 35*(5), 497–502.
21. Lynge, E., Sandegaard, J. L., & Rebolj, M. (2011). The Danish national patient register. *Scandinavian Journal of Public Health, 39*(7 Suppl), 30–33.
22. Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases, 40*(5), 373–383.
23. Osborne, R. H., Batterham, R. W., Elsworth, G. R., Hawkins, M., & Buchbinder, R. (2013). The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC public health, 13,* 658-2458-13-658.
24. Maindal, H. T., Kayser, L., Norgaard, O., Bo, A., Elsworth, G. R., & Osborne, R. H. (2016). Cultural adaptation and validation of the Health Literacy Questionnaire (HLQ): Robust nine-dimension Danish language confirmatory factor model. *SpringerPlus, 5*(1), 1232-016-2887-9. **eCollection 2016**.
25. Luszczynska, A., Scholz, U., & Schwarzer, R. (2005). The general self-efficacy scale: Multicultural validation studies. *The Journal of Psychology, 139*(5), 439–457.
26. Schwarzer R, & Jerusalem M. (1995). Generalized self-efficacy scale. In: J. Weinman, S.Wright, & M. Johnston, editors. Measures in health psychology: A user's portfolio: Causal and control beliefs. Windsor, England: NFER-NELSON; 1995. pp. 35–7.
27. Scholz, U., Gutiérrez-Doña, B., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment, 18*(3), 242–251.
28. Bech, P., Gudex, C., & Johansen, K. S. (1996). The WHO (Ten) well-being index: Validation in diabetes. *Psychotherapy and Psychosomatics, 65*(4), 183–190.
29. Topp, C. W., Ostergaard, S. D., Sondergaard, S., & Bech, P. (2015). The WHO-5 well-being index: A systematic review of the literature. *Psychotherapy and Psychosomatics, 84*(3), 167–176.
30. Bjorner, J. B., Damsgaard, M. T., Watt, T., & Groenvold, M. (1998). Tests of data quality, scaling assumptions, and reliability of the Danish SF-36. *Journal of Clinical Epidemiology, 51*(11), 1001–1011.
31. Bjorner, J. B., Thunedborg, K., Kristensen, T. S., Modvig, J., & Bech, P. (1998). The Danish SF-36 Health Survey: Translation and preliminary validity studies. *Journal of Clinical Epidemiology, 51*(11), 991–999.
32. Maindal, H. T., Sokolowski, I., & Vedsted, P. (2009). Translation, adaptation and validation of the American short form Patient Activation Measure (PAM13) in a Danish version. *BMC Public Health, 9,* 209-2458-9-209.
33. Overgaard, M., Andersen, P. K., & Parner, E. T. (2015). Regression analysis of censored data using pseudo-observations: An update. *Stata Journal, 15*(3), 809–821.
34. Klein, J. P., Logan, B., Harhoff, M., & Andersen, P. K. (2007). Analyzing survival curves at a fixed point in time. *Statistics in Medicine, 26*(24), 4505–4519.
35. Biering, K., Hjollund, N. H., & Frydenberg, M. (2015). Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. *Clinical Epidemiology, 7,* 91–106.
36. Barlow, J., Wright, C., Sheasby, J., Turner, A., & Hainsworth, J. (2002). Self-management approaches for people with chronic conditions: A review. *Patient Education and Counseling, 48*(2), 177–187.
37. Wagner, E. H., Austin, B. T., Davis, C., Hindmarsh, M., Schaefer, J., & Bonomi, A. (2001). Improving chronic illness care: translating evidence into action. *Health affairs (Project Hope), 20*(6), 64–78.
38. Abu Dabrh, A. M., Gallacher, K., Boehmer, K. R., Hargraves, I. G., & Mair, F. S. (2015). Minimally disruptive medicine: the evidence and conceptual progress supporting a new era of healthcare. *The journal of the Royal College of Physicians of Edinburgh, 45*(2), 114–117.
39. John, K., Tailor, S., Anderson, J., & Lawthom, C. (2019). Managing epilepsy in austerity—Evaluating the utility and value of the epilepsy specialist nurse in an open access model of service delivery: Aneurin Bevan Epilepsy Specialist Team (A.B.E.S.T.). *Seizure, 65,* 98–100.
40. May, C., Montori, V. M., & Mair, F. S. (2009). We need minimally disruptive medicine. *BMJ (Clinical Research Ed.), 339,* b2803.
41. Jacoby, A., & Baker, G. A. (2008). Quality-of-life trajectories in epilepsy: A review of the literature. *Epilepsy & Behavior: E&B, 12*(4), 557–571.
42. Josephson, C. B., Patten, S. B., Bulloch, A., Williams, J. V. A., Lavorato, D., Fiest, K. M., et al. (2017). The impact of seizures

on epilepsy outcomes: A national, community-based survey. *Epilepsia, 58*(5), 764–771.
43. Mejdahl, C. T., Schougaard, L. M. V., Hjollund, N. H., Riiskjaer, E., & Lomborg, K. (2018). Exploring organisational mechanisms in PRO-based follow-up in routine outpatient care—an interpretive description of the clinician perspective. *BMC health services research, 18*(1), 546-018-3352-y.
44. Jeppesen, M. M., Mogensen, O., Hansen, D. G., Iachina, M., Korsholm, M., & Jensen, P. T. (2017). Detection of recurrence in early stage endometrial cancer—the role of symptoms and routine follow-up. *Acta Oncologica (Stockholm, Sweden), 56*(2), 262–269.
45. Tecson, K. M., Bass, K., Felius, J., Hall, S. A., Jamil, A. K., & Carey, S. A. (2019). Patient "activation" of patients referred for advanced heart failure therapy. *The American Journal of Cardiology, 123*(4), 627–631.
46. Mejdahl, C. T., Schougaard, L. M. V., Hjollund, N. H., Riiskjaer, E., Thorne, S., & Lomborg, K. (2018). PRO-based follow-up as a means of self-management support—an interpretive description of the patient perspective. *Journal of Patient-Reported Outcomes, 2,* 38-018-0067-0. **eCollection 2017**.
47. Aaby, A., Friis, K., Christensen, B., Rowlands, G., & Maindal, H. T. (2017). Health literacy is associated with health behaviour and self-reported health: A large population-based study in individuals with cardiovascular disease. *European Journal of Preventive Cardiology, 24*(17), 1880–1888.
48. Chahardah-Cherik, S. M. S., Gheibizadeh, M. P., Jahani, S. P., & Cheraghian, B. P. (2018). The relationship between health literacy and health promoting behaviors in patients with type 2 diabetes. *International Journal of Community Based Nursing and Midwifery, 6*(1), 65–75.
49. Olesen, K., Reynheim, A. L. F., Joensen, L., Ridderstrale, M., Kayser, L., Maindal, H., et al. (2017). Higher health literacy is associated with better glycemic control in adults with type 1 diabetes: A cohort study among 1399 Danes. *BMJ Open Diabetes Research & Care, 5*(1), e000437.
50. Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ (Clinical Research Ed.), 332*(7549), 1080.
51. Sundheds- og Ældreministeriet. Bekendtgørelse af lov om videnskabsetisk behandling af sundhedsvidenskabelige forskningsprojekter [in Danish]. Retrieved September 23 2019 from https://www.retsinformation.dk/Forms/r0710.aspx?id=192671#idfae75201-60d1-46a6-bef7-ccd4ce8f39aa.
52. Datatilsynet. Vejledning Samtykke [in Danish]. Retrieved September 23 2019 from https://www.datatilsynet.dk/media/6562/samtykke.pdf.

Affiliations

Liv Marit Valen Schougaard[1] · Annette de Thurah[2,3] · Jakob Christensen[3,4,5] · Kirsten Lomborg[3] · Helle Terkildsen Maindal[6] · Caroline Trillingsgaard Mejdahl[1] · Jesper Medom Vestergaard[7] · Trine Nøhr Winding[7] · Karin Biering[3,7] · Niels Henrik Hjollund[1,8]

1 AmbuFlex/WestChronic, Occupational Medicine, University Research Clinic, Aarhus University, Gl. Landevej 61, 7400 Herning, Denmark

2 Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark

3 Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

4 Department of Neurology, Aarhus University Hospital, Aarhus, Denmark

5 National Centre for Register-based Research, Department of Economics and Business Economics, Aarhus BSS, Aarhus University, Aarhus, Denmark

6 Department of Public Health, Aarhus University, Aarhus, Denmark

7 Department of Occupational Medicine, University Research Clinic, Regional Hospital West Jutland, Herning, Denmark

8 Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

TITLE: Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study

AUTHORS: Liv Marit Valen Schougaard[1], Annette de Thurah[2,3], Jakob Christensen[3,4,5], Kirsten Lomborg[3], Helle Terkildsen Maindal[6], Caroline Trillingsgaard Mejdahl[1], Jesper Medom Vestergaard[7], Trine Nøhr Winding[7], Karin Biering[3,7], Niels Henrik Hjollund[1,8]

AFFILIATIONS:

[1]AmbuFlex/WestChronic, Occupational Medicine, University Research Clinic, Aarhus University, Herning, Denmark

[2] Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark

[3] Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

[4] Department of Neurology, Aarhus University Hospital, Aarhus, Denmark

[5] National Centre for Register-based Research, Department of Economics and Business Economics, Aarhus BSS, Aarhus University, Denmark

[6] Department of Public Health, Aarhus University, Aarhus, Denmark

[7] Department of Occupational Medicine, University Research Clinic, Regional Hospital West Jutland, Herning, Denmark

[8] Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

CORRESPONDING AUTHOR: Liv Marit Valen Schougaard, email: livschou@rm.dk

Appendix 1: Multiple imputation models

Based on the assumption that data were missing at random, 100 complete datasets were created based on a model (model 1) of all relevant variables measured in the population (age, gender, cohabitation status, education, household income, labour market affiliation, co-morbidity, psychiatric diseases, duration of epilepsy diagnosis, and questionnaire scores). The robustness of the imputed model was evaluated by modifying the variables in the model (model 2 and 3).

	MODEL 1																								
(	logit	, include(	alder	sex		udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode		psyk		hlq_4_score	hlq_6_score	hlq_9_score	gse_score					pseudo1	pseudo2	pseudo3	))	bor_alene	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode		psyk				hlq_9_score	gse_score					pseudo1	pseudo2	pseudo3	))	AEKVIVADISP_13_grp_recode	///
(	ologit	, include(	alder	sex	bor_alene		AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk		hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	udd_grp	///
(	regress	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode		psyk			hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	hlq_4_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode	charlson_index_lpr_recode	psyk	dg40_varighed_aar_ny	hlq_4_score		hlq_9_score	gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	hlq_6_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk	dg40_varighed_aar_ny	hlq_4_score	hlq_6_score		gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	hlq_9_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk	dg40_varighed_aar_ny	hlq_4_score	hlq_6_score	hlq_9_score		who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	gse_score	///
(	regress	, include(	alder	sex				dream_grp_recode	charlson_index_lpr_recode	psyk		hlq_4_score	hlq_6_score	hlq_9_score	gse_score		sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	who5	///
(	ologit	, include(	alder	sex			AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk		hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5		pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	sf_gh1h	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk		hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h		pam_12	pseudo1	pseudo2	pseudo3	))	pam_5	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk		hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h	pam_5		pseudo1	pseudo2	pseudo3	))	pam_12	///
	MODEL 2																								
(	logit	, include(	alder	sex		udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode				hlq_4_score	hlq_6_score	hlq_9_score	gse_score					pseudo1	pseudo2	pseudo3	))	bor_alene	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode						hlq_9_score	gse_score					pseudo1	pseudo2	pseudo3	))	AEKVIVADISP_13_grp_recode	///
(	ologit	, include(	alder	sex	bor_alene		AEKVIVADISP_13_grp_recode	dream_grp_recode				hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	udd_grp	///
(	regress	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode					hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	hlq_4_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode				hlq_4_score		hlq_9_score	gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	hlq_6_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode				hlq_4_score	hlq_6_score		gse_score	who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	hlq_9_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode				hlq_4_score	hlq_6_score	hlq_9_score		who5	sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	gse_score	///
(	regress	, include(	alder	sex				dream_grp_recode				hlq_4_score	hlq_6_score	hlq_9_score	gse_score		sf_gh1h	pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	who5	///
(	ologit	, include(	alder	sex			AEKVIVADISP_13_grp_recode	dream_grp_recode				hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5		pam_5	pam_12	pseudo1	pseudo2	pseudo3	))	sf_gh1h	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode				hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h		pam_12	pseudo1	pseudo2	pseudo3	))	pam_5	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode				hlq_4_score	hlq_6_score	hlq_9_score	gse_score	who5	sf_gh1h	pam_5		pseudo1	pseudo2	pseudo3	))	pam_12	///
	MODEL 3																								
(	logit	, include(	alder	sex		udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode		psyk										pseudo1	pseudo2	pseudo3	))	bor_alene	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode		psyk										pseudo1	pseudo2	pseudo3	))	AEKVIVADISP_13_grp_recode	///
(	ologit	, include(	alder	sex	bor_alene		AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk										pseudo1	pseudo2	pseudo3	))	udd_grp	///
(	regress	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode		psyk										pseudo1	pseudo2	pseudo3	))	hlq_4_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp		dream_grp_recode	charlson_index_lpr_recode	psyk	dg40_varighed_aar_ny									pseudo1	pseudo2	pseudo3	))	hlq_6_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk	dg40_varighed_aar_ny									pseudo1	pseudo2	pseudo3	))	hlq_9_score	///
(	regress	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk	dg40_varighed_aar_ny									pseudo1	pseudo2	pseudo3	))	gse_score	///
(	regress	, include(	alder	sex				dream_grp_recode	charlson_index_lpr_recode	psyk										pseudo1	pseudo2	pseudo3	))	who5	///
(	ologit	, include(	alder	sex			AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk										pseudo1	pseudo2	pseudo3	))	sf_gh1h	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk										pseudo1	pseudo2	pseudo3	))	pam_5	///
(	ologit	, include(	alder	sex	bor_alene	udd_grp	AEKVIVADISP_13_grp_recode	dream_grp_recode	charlson_index_lpr_recode	psyk										pseudo1	pseudo2	pseudo3	))	pam_12	///

Appendix 2: Original raw analyses

Risk ratio (RR) of referral to PRO-based follow-up 6, 12, and 18 months after the first visit at Department of Neurology, Aarhus University Hospital according to register determinants (*N*=802)

	6-month follow-up		12-month follow-up		18-month follow-up	
	Crude RR	**Adjusted RR[a]**	**Crude RR**	**Adjusted RR[a]**	**Crude RR**	**Adjusted RR[a]**
Age, years						
15−24	Ref	Ref	Ref	Ref	Ref	Ref
25−39	1.03 (0.63−1.67)	0.98 (0.58−1.66)	0.90 (0.59−1.37)	0.84 (0.53−1.31)	0.89 (0.59−1.32)	0.79 (0.51−1.24)
40−59	1.04 (0.66−1.65)	0.87 (0.52−1.44)	0.98 (0.67−1.44)	0.83 (0.54−1.28)	0.97 (0.67−1.40)	0.83 (0.55−1.25)
60−69	1.08 (0.65−1.77)	0.74 (0.42−1.30)	0.97 (0.63−1.48)	0.74 (0.45−1.21)	0.83 (0.55−1.26)	0.66 (0.40−1.06)
70−99	0.97 (0.61−1.54)	0.85 (0.48−1.49)	0.85 (0.57−1.27)	0.95 (0.57−1.61)	0.78 (0.53−1.16)	0.87 (0.52−1.43)
Gender						
Female	Ref	Ref	Ref	Ref	Ref	Ref
Male	1.32 (0.97−1.79)	1.33 (0.94−1.87)	1.29 (0.99−1.69)	1.28 (0.97−1.70)	1.39 (1.07−1.80)	1.34 (1.02−1.75)
Cohabitation status						
Living with a partner/family	Ref	Ref	Ref	Ref	Ref	Ref
Living alone	0.62 (0.44−0.87)	0.65 (0.44−0.95)	0.56 (0.41−0.76)	0.61 (0.43−0.86)	0.58 (0.43−0.79)	0.67 (0.48−0.93)
Education						
High (> 12 years)	Ref	Ref	Ref	Ref	Ref	Ref
Medium (10−12 years)	1.06 (0.72−1.55)	1.00 (0.67−1.49)	1.31 (0.91−1.87)	1.25 (0.87−1.78)	1.26 (0.88−1.78)	1.19 (0.85−1.68)
Low (< 10 years)	0.56 (0.36−0.86)	0.47 (0.29−0.76)	0.69 (0.47−1.03)	0.63 (0.41−0.96)	0.73 (0.50−1.07)	0.66 (0.44−0.98)
Household income						
High	Ref	Ref	Ref	Ref	Ref	Ref
Medium	0.46 (0.31−0.66)	0.61 (0.39−0.96)	0.52 (0.38−0.71)	0.71 (0.48−1.04)	0.49 (0.36−0.67)	0.67 (0.46−0.97)
Low	0.50 (0.35−0.72)	0.65 (0.43−0.99)	0.44 (0.32−0.62)	0.59 (0.41−0.85)	0.43 (0.31−0.59)	0.52 (0.36−0.75)
Labour market affiliation						
Self-supporting	Ref	Ref	Ref	Ref	Ref	Ref
Normal retirement	0.69 (0.48−1.00)	0.71 (0.43−1.19)	0.71 (0.52−0.98)	0.94 (0.57−1.54)	0.63 (0.46−0.86)	0.82 (0.51−1.34)
Temporary social benefits	0.69 (0.46−1.02)	0.68 (0.45−1.03)	0.76 (0.54−1.07)	0.79 (0.55−1.11)	0.67 (0.49−0.94)	0.66 (0.47−0.93)
Permanent social benefits	0.38 (0.21−0.66)	0.52 (0.27−0.99)	0.43 (0.26−0.70)	0.56 (0.32−0.98)	0.39 (0.24−0.62)	0.48 (0.28−0.82)
Co-morbidity (Charlson Index)						
Low 0	Ref	Ref	Ref	Ref	Ref	Ref
Medium 1−2	0.77 (0.53−1.12)	0.73 (0.48−1.11)	0.70 (0.50−0.97)	0.67 (0.46−0.97)	0.71 (0.52−0.97)	0.73 (0.51−1.03)
High > 2	1.02 (0.66−1.59)	0.98 (0.61−1.57)	0.79 (0.52−1.20)	0.80 (0.51−1.24)	0.68 (0.44−1.05)	0.72 (0.45−1.13)

Psychiatric disease						
No	Ref	Ref	Ref	Ref	Ref	Ref
Yes	0.65 (0.37−1.15)	0.77 (0.43−1.38)	0.48 (0.27−0.86)	0.57 (0.31−1.04)	0.46 (0.25−0.83)	0.52 (0.28−0.96)

Numbers in round brackets are 95% confidence intervals (CIs).The estimated RRs and 95% CIs were obtained after multiple imputations in a generalised linear regression using the pseudo-value approach.

[a] Mutual adjusted for age, gender, cohabitation status, education, and co-morbidity

Risk ratio (RR) of referral to PRO-based follow-up 6, 12, and 18 months after the first visit at Department of Neurology, Aarhus University Hospital according to questionnaire determinants (*N*=411)

	6-month follow-up		**12-month follow-up**		**18-month follow-up**	
	Crude RR	**Adjusted RR[a]**	**Crude RR**	**Adjusted RR[a]**	**Crude RR**	**Adjusted RR[a]**
Social support for health (HLQ4)						
High (> 2)	Ref	Ref	Ref	Ref	Ref	Ref
Low (≤ 2)	0.57 (0.13−2.58)	0.32 (0.04−2.45)	0.50 (0.12−2.03)	0.32 (0.04−2.32)	0.45 (0.10−1.93)	0.34 (0.05−2.51)
Ability to actively engage with healthcare providers (HLQ6)						
High (> 3)	Ref	Ref	Ref	Ref	Ref	Ref
Low (≤ 3)	0.70 (0.42−1.19)	0.78 (0.42−1.45)	0.64 (0.39−1.04)	0.68 (0.38−1.23)	0.64 (0.39−1.05)	0.69 (0.39−1.23)
Understanding health information well enough to know what to do (HLQ9)						
High (> 3)	Ref	Ref	Ref	Ref	Ref	Ref
Low (≤ 3)	0.41 (0.21−0.82)	0.48 (0.23−0.98)	0.32 (0.16−0.65)	0.35 (0.17−0.74)	0.33 (0.17−0.66)	0.35 (0.17−0.73)
Self-efficacy (GSE)						
High (≥ 30)	Ref	Ref	Ref	Ref	Ref	Ref
Low (< 30)	0.63 (0.43−0.93)	0.73 (0.48−1.13)	0.64 (0.45−0.90)	0.72 (0.49−1.07)	0.64 (0.46−0.90)	0.70 (0.49−1.01)
Well-being (WHO-5)						
High (≥ 50)	Ref	Ref	Ref	Ref	Ref	Ref
Low (< 50)	0.71 (0.45−1.12)	0.78 (0.46−1.30)	0.57 (0.37−0.89)	0.62 (0.38−1.02)	0.50 (0.32−0.79)	0.57 (0.36−0.92)
General health						
Excellent/ Very good	Ref	Ref	Ref	Ref	Ref	Ref
Good	1.21 (0.80−1.82)	1.20 (0.76−1.87)	0.99 (0.70−1.41)	0.97 (0.63−1.48)	0.87 (0.62−1.21)	0.84 (0.57−1.23)
Fair/ Poor	0.66 (0.39−1.14)	0.73 (0.42−1.27)	0.53 (0.32−0.86)	0.61 (0.36−1.02)	0.43 (0.26−0.71)	0.48 (0.29−0.82)
Patient activation [b]						
Agree Strongly/ Agree	Ref	Ref	Ref	Ref	Ref	Ref
Disagree Strongly/ Disagree	0.54 (0.33−0.87)	0.54 (0.29−1.02)	0.57 (0.37−0.86)	0.59 (0.35−0.99)	0.49 (0.32−0.76)	0.50 (0.30−0.85)
Patient activation [c]						
Agree Strongly/ Agree	Ref	Ref	Ref	Ref	Ref	Ref
Disagree Strongly/ Disagree	0.53 (0.33−0.87)	0.51 (0.29−0.91)	0.52 (0.34−0.81)	0.52 (0.31−0.89)	0.46 (0.29−0.72)	0.45 (0.26−0.77)

Abbreviations HLQ: Health Literacy Questionnaire; GSE: General Self-efficacy scale; WHO-5: WHO-Five Well-being Index

Numbers in round brackets are 95% confidence intervals (CIs).The estimated RRs and 95% CIs were obtained after multiple imputations in a generalised linear regression using the pseudo-value approach. [a] Adjusted for age, gender, cohabitation status, education, and co-morbidity

[b] I am confident that I can tell when I need to get outpatient care

[c] I am confident I can figure out solutions when new situations or problems arise with my health condition

Appendix 3: Sensitivity analysis

We assumed that self-reported health literacy missing data were lower than expected from the imputed dataset. For patients with missing self-reported health literacy data, health literacy scores were reduced with one point corresponding to approximately one standard deviation. Subsequently, the cumulative risk ratio (RR) was analysed by a generalised linear regression using the pseudo-value approach at the three time points.

Risk ratio (RR) of referral to PRO-based follow-up 6, 12, and 18 months after the first visit at Department of Neurology, Aarhus University Hospital according to self-reported health literacy (*N*=802)

	6-month follow-up		**12-month follow-up**		**18-month follow-up**	
	Crude RR	**Adjusted RR[a]**	**Crude RR**	**Adjusted RR[a]**	**Crude RR**	**Adjusted RR[a]**
Social support for health (HLQ4 score)	1.45 (1.19−1.77)	1.32 (1.04−1.67)	1.38 (1.16−1.65)	1.24 (1.02−1.51)	1.42 (1.18−1.71)	1.30 (1.06−1.59)
Ability to actively engage with healthcare providers (HLQ6 score)	1.40 (1.19−1.64)	1.27 (1.04−1.55)	1.38 (1.20−1.58)	1.28 (1.10−1.49)	1.39 (1.21−1.58)	1.30 (1.13−1.51)
Understanding health information well enough to know what to do (HLQ9 score)	1.42 (1.22−1.65)	1.30 (1.08−1.56)	1.35 (1.19−1.54)	1.26 (1.09−1.46)	1.38 (1.21−1.57)	1.31 (1.13−1.52)

Abbreviations HLQ: Health Literacy Questionnaire; GSE: General Self-efficacy scale; WHO-5: WHO-Five Well-being Index
Numbers in round brackets are 95% confidence intervals (CIs).The estimated RRs and 95% CIs were obtained after multiple imputations in a generalised linear regression using the pseudo-value approach.
[a] Adjusted for age, gender, cohabitation status, education, and co-morbidity

PAPER II

Including online supplemental materials:

Appendix 1: The questionnaire and PRO-algorithm development process

Table 1: Agreement and reliability between the items from test 1 to test 2 in original categories and categories within the framework of the PRO- algorithm

Note: The epilepsy questionnaire can be found in the Appendices section (Appendix II)

BMJ Open

Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test–retest reliability study

Liv Marit Valen Schougaard,[1] Annette de Thurah,[2,3] David Høyrup Christiansen,[4] Per Sidenius,[5] Niels Henrik Hjollund[1,6]

To cite: Schougaard LMV, de Thurah A, Christiansen DH, *et al*. Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test–retest reliability study. *BMJ Open* 2018;**8**:e021337. doi:10.1136/bmjopen-2017-021337

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-021337).

PS is deceased

Received 22 December 2017
Revised 26 March 2018
Accepted 6 June 2018

For numbered affiliations see end of article.

Correspondence to
Dr Liv Marit Valen Schougaard; livschou@rm.dk

ABSTRACT

Objectives Patient-reported outcome (PRO) measures have been used in epilepsy outpatient clinics in Denmark since 2011. The patients' self-reported PRO data are used by clinicians as a decision aid to support whether a patient needs contact with the outpatient clinic or not based on a PRO algorithm. Validity and reliability are fundamental to any PRO measurement used at the individual level in clinical practice. The aim of this study was to evaluate the test–retest reliability of the PRO algorithm used in epilepsy outpatient clinics and to analyse whether the method of administration (web and paper) would influence the result.

Design and setting Test–retest reliability study conducted in three epilepsy outpatient clinics in Central Denmark Region, Denmark.

Participants A total of 554 epilepsy outpatients aged 15 years or more were included from August 2016 to April 2017. The participants completed questionnaires at two time points and were randomly divided into four test–retest groups: web–web, paper–paper, web–paper and paper–web. In total, 166 patients completed web–web, 112 paper–paper, 239 web–paper and 37 paper–web.

Results Weighted kappa with squared weight was 0.67 (95% CI 0.60 to 0.74) for the pooled PRO algorithm, and perfect agreement was observed in 82% (95% CI 78% to 85%) of the cases. There was a tendency towards higher test–retest reliability and agreement estimates within same method of administration (web–web or paper–paper) compared with a mixture of methods (web–paper and paper–web).

Conclusions The PRO algorithm used for clinical decision support in epilepsy outpatient clinics showed moderate to substantial test–retest reliability. Different methods of administration produced similar results, but an influence of change in administration method cannot be ruled out.

Strengths and limitations of this study

- This study explores the quality in terms of test–retest reliability of a patient-reported outcome instrument used as a decision aid for identifying outpatients in need of clinical attention.
- The study contributes with knowledge whether the method of administration (web, paper or a mixture of the two modalities) influences the results.
- The study includes a large sample size, however, the response rate was low.
- The study population was a homogeneous and healthier group of patients compared with the non-responders, which may have lead to underestimation of the results.
- The study has low prevalence of the measured event and this could affect the agreement estimates.

INTRODUCTION

Patient-reported outcome (PRO) measures are defined as a measurement concerning the patient's health status reported directly from the patient.[1] The use of PRO measures in clinical practice has increased during the last decade, and potential benefits have been described such as better patient–clinician communication, better identification of patients' functional or mental health issues, better monitoring of treatment on patients' health, a better tool to inform clinical decision-making and support patient self-management.[2–5] However, barriers have been identified as well, for example, practising physicians prefer talking to the patients rather than using standardised PRO measures.[6] Furthermore, if clinicians do not rely on the PRO measures to judge treatment, the use of PRO may raise concerns related to both validity and interpretation.[7 8] PRO measures are typically developed for research purposes and used at an aggregated level.[9] These measures are not necessarily suitable for use in clinical practice. PRO measures used in clinical practice at the individual level should reflect clinically relevant aspects and should be meaningful to patients as well as clinicians.[10] Furthermore, validity and reliability are

fundamental to any PRO measurement used at the aggregated level in research as well as at the individual level in clinical practice.[11]

Epilepsy is a long-term chronic condition affecting approximately 1% of the general population.[12] Epilepsy represents a major socioeconomic burden for patients as well as for society.[13] The condition is characterised by recurrent seizures affecting physiological, psychological and social aspects of daily life,[14 15] aspects that can only be reported by the patients themselves. However, PROs have not been routinely collected in neurological outpatient clinics. A study that included patients with epilepsy as well as other neurological conditions concluded that systematic collection of PROs may be feasible in a clinical setting.[16] Additional studies regarding use of PROs in epilepsy clinics have not been identified, but the way epilepsy is managed differs greatly between countries.[17]

In Denmark, outpatient follow-up in patients with epilepsy has traditionally been based on regular consultations at a neurological department. However, since 2011, PROs have been used in three epilepsy outpatient clinics in the Central Denmark Region.[18] The clinicians use PRO measures as the basis for outpatient follow-up. Instead of prescheduled appointments, the patients fill in either a web or paper questionnaire at home regarding daily life with epilepsy. The patients' self-reported PRO data are used by clinicians as a decision aid to support whether a patient needs contact with the clinic or not based on an automated PRO algorithm.[18] Furthermore, the PRO data are used to monitor treatment effects and potential side effects, and to facilitate patient-centred communication between the patient and the clinician.[18] As of October 2017, approximately 5000 outpatients have been referred to PRO-based follow-up in three epilepsy outpatient clinics in Central Denmark Region. The Danish government and the regions, who run the public hospitals, have decided on a national strategy regarding implementation of PROs in patients with epilepsy before 2020.

In 2011, a disease-specific PRO instrument combined with a PRO algorithm used as decision aid in outpatients with epilepsy was developed and tested in close cooperation with clinicians and patients from three epilepsy outpatient clinics in Denmark. Content and face validity have been crucial during the development process. The test–retest reliability of the PRO algorithm and the questionnaire has not been evaluated, but is pivotal in the development of the instrument.[19] Furthermore, few test–retest studies[20 21] have evaluated whether the method of administration has any influence on the results.

AIMS

The aim of this study was to evaluate the test–retest reliability of the PRO algorithm used for clinical decision support in epilepsy outpatient follow-up and to analyse to what extent the four different methods of administration (web–web, paper–paper, web–paper and paper–web) would influence the result. A further aim was to evaluate the test–retest reliability of the single items included in the questionnaire.

METHODS

The epilepsy questionnaire

Development

Clinicians working with epilepsy experienced an increased volume of patients in the outpatient clinic and the majority of these patients were well treated. However, the need of monitoring treatment effect and screen for functional and mental health issues were still necessary. Therefore, self-reported data collected from the patients' home were assumed to have a great potential in this patient group. Several epilepsy-specific PRO instruments have been developed[22]; however, no established instruments covering the purpose of identifying patients who need clinical attention were found. In 2011, a research consensus team that included clinical experts and experts in PRO provided inputs to the content and construct of an epilepsy questionnaire. The purpose was to develop an instrument which could screen for epilepsy patients' health problems to support clinical decision-making in outpatient follow-up.[10 18] The target group was patients with epilepsy ≥15 years with no cognitive impairments. The content was based on validated PRO instruments or items; however, ad hoc items were developed if existing instruments or items were not available. This process was based on inputs from specialists in epilepsy, a literature search and interviews with patients.[23] The first version of the questionnaire was pretested by using semistructural interviewing techniques in 20 representative epilepsy patients from two outpatient clinics in Central Denmark Region. The aim of the pilot test was to identify potential problems such as low relevance of items, ambiguity of items and lack of important topics.[24] The majority of the patients found the questionnaire content relevant, and no critical comprehension difficulties were identified. Some patients pointed out recall problems regarding some of the seizure items. They did not report lack of any essential topics nor did the time used to fill in the questionnaire raise any criticism. Subsequently, the PRO questionnaire was implemented and used in clinical practice, and experiences have been evaluated yearly since 2011 at consensus meetings.[18] Additionally, information regarding the development process and the fourth version of the questionnaire can be found in the online supplementary material.

Content

The questionnaire included information specific to aspects of daily life with epilepsy, for example, seizures, side effects, well-being, general health and social problems. The questionnaire included WHO-5 Well-Being Index (WHO-5),[25] items from the Short-Form 36 (SF-36)[26] and items from the Symptom Checklist 92 (SCL-92).[27] WHO-5 is a generic questionnaire including five items reflecting the construct mental well-being.[28] The instrument has

demonstrated sufficient psychometric properties in other patient populations.[28] The percentages scores range from 0 to 100, and a percentage score below 50 indicates increased risk of poor mental well-being, and an evaluation for depression is recommended. SF-36 is a generic questionnaire with eight subscales measuring physical and mental health,[26] and the psychometric properties of the Danish SF-36 have been documented.[29] Two single items regarding general health from SF-36 were included in the epilepsy questionnaire. SCL-92 consists of nine subscales measuring, for example, somatisation, anxiety and depression, and validity has previously been measured in a Danish population.[27] Ten single items from SCL-92 have been used in the epilepsy questionnaire, three of which have been partly modified. In addition, the epilepsy questionnaire included self-composed items, for example, regarding seizures, symptoms, medication adherence and pregnancy. Online supplementary appendix 1 presents the items evaluated in this study.

Decision aid

The questionnaire is used to support clinical decision-making in clinical practice. A clinical expert group in epilepsy has assigned the response options for each item in three colours: green, yellow or red based on what the doctors considered clinically important to react on to identify patients with need of attention. The colours represent a computerised algorithm, which is processed automatically by AmbuFlex's web server,[10] for example, if only one item response category was red, the whole response was given a red colour. A red colour indicates that the patient needs or wishes contact with the outpatient clinic, whereas a yellow colour indicates that the patient may need contact with the clinic. An overview of the response is embedded in the electronic health record (EHR). In yellow cases, a clinician assesses the overview, and based on the PRO data and other information in the patient's EHR, it is decided whether further contact is needed. A green colour indicates that the patient does not need or wish contact with the clinic, and a subsequent questionnaire is sent to the patient at a predefined interval (eg, after 3, 6 or 12 months). A patient can overrule a green and yellow algorithm by the item 'What is your present need for contact with the outpatient clinic.' By such a request, the whole response will always turn red. This item was not included in the retest study since this statement would probably change from test 1 to test 2 due to action taken based on PRO data in test 1, thus indicating responsiveness rather than reliability. Online supplementary appendix 1 presents an overview of the red, yellow and green item response categories evaluated in this study.

Patient and public involvement

A total of 20 patients were involved in the development process of the questionnaire. They have contributed with valuable insight to both face and content validity. Furthermore, feedback from patients after implementation has been included during a yearly questionnaire revision. Patients were not involved in the design, recruitment or conduct of this study.

Study population and procedure

Outpatients with epilepsy aged 15 years or more and referred to PRO-based follow-up from the three epilepsy outpatient clinics in Central Denmark Region were included. Data collection took place from August 2016 to April 2017. The general recommendation regarding sample size in reliability studies is to include at least 50 participants.[30] In this study, an increased number of patients were included due to an expected risk of low prevalence in some items and further to gain the opportunity to conduct subanalyses with different test–retest patterns. The participants completed questionnaires at two time points. First, they responded to the normal prescheduled epilepsy questionnaire from the outpatient clinics as planned (named test 1). Patients referred to PRO-based follow-up can select which administration method they prefer, although the web-based method is recommended. In the present study, participants answered test 1 by their preferred method. Subsequently, a letter was sent to the participants who were asked to complete the same questionnaire after approximately 2 weeks (named test 2). According to experiences with the Danish postal service in other WestChronic projects,[10] the date of dispatch of the letter was different in web and paper responders. The letter was sent 8 days after received date of the questionnaire in test 1 in web responders and after 4 days in paper responders. No reminders were sent in test 2. Participants were randomly divided into groups with four test–retest patterns: web–web, paper–paper, web–paper and paper–web. From August 2016 to November 2016, the randomisation allocation was 1:1 in both paper and web responders. Due to a low response rate in the paper–web group, the allocation was changed for paper responders. From the end of November 2016 to April 2017, the randomisation allocation was 0.25 in the paper–paper group and 0.75 in the paper–web group.

Data analysis

In nominal and ordinal data, respectively, unweighted and weighted kappa statistics with squared weights were used to assess reliability.[19] The 95% CIs for weighted kappa values were measured using non-parametric bootstrap methods (1000 replications).[31] The kappa values were interpreted as follows: <0.2 (slight), 0.21–0.4 (fair), 0.41–0.6 (moderate), 0.61–0.8 (substantial) and 0.81–1.0 (almost perfect).[32] Proportion of agreement was used to assess agreement measures.[19] Due to a small number of participants in the paper–web group, the two mixed groups (web–paper and paper–web) were merged in the analyses. A sensitivity analysis with a shorter time interval was estimated for both the PRO algorithm and for the different modes of administration by excluding participants with intervals above the median number of days between test 1 and test 2. The interval between test 1 and

test 2 was calculated as the difference in number of days from the date of response. In paper responses, the interval was calculated as the date of received questionnaires minus 4 days. For example, the received response date 10 October became 6 October. This decision was made based on experiences with the postal service in other WestChronic projects.[10] Differences between responders and non-responders at test 2 were evaluated by X^2 test for categorical variables or the Kruskal-Wallis test for continuous variables based on data from test 1.

Test–retest reliability and agreement were assessed both within the item categories and according to the red, yellow or green item algorithm categories. For example, the item concerning headaches was assessed at two and five levels. The five levels were the original scale 'never', 'occasionally', 'sometimes', 'often' and 'very often', whereas the two levels were according to the predefined PRO algorithm and in this case green or yellow. 'Never', 'occasionally' and 'sometimes' were grouped into green, and 'often' and 'very often' were grouped into yellow. Lack of response was assessed for all items and was considered not acceptable if data were missing in more than 5% of an item category. Floor and ceiling effects were assessed and considered present if a high proportion (more than 15%) of the respondents had a score at the lower or upper end of the scale.[33]

Table 1 Patient characteristic measured in test 1 in responders and non-responders in test 2 among outpatients with epilepsy, n=1640

	Responders (n=554) n (%)	Non-responders (n=1086) n (%)
Gender, men	286 (52)	511 (47)
Age, year, median (IQR)	57.3 (42.7 to 67.7)	49.7 (33.8 to 64.8)
Department		
Aarhus	409 (74)	831 (77)
Holstebro	115 (21)	174 (16)
Viborg	30 (5)	81 (7)
Patient-reported outcome algorithm in test 1		
Green	116 (21)	200 (18)
Yellow	349 (63)	670 (62)
Red	89 (16)	216 (20)
WHO-5 Well-Being Index, median (IQR)	76 (60 to 84)	72 (56 to 80)
General health		
Excellent/very good	258 (47)	448 (41)
Good	209 (38)	427 (39)
Fair/poor	87 (16)	206 (19)
Missing item categories		5 (1)

RESULTS

Patient characteristics

A total of 554/1640 participants responded to the questionnaire in test 2, corresponding to a response rate of 34%. The median age was 57.3 years, with an IQR of 42.7 to 67.7 years. Non-responders in test 2 were more likely younger (p<0.001), paper–responders in test 1 (p<0.001) had lower self-reported well-being (p=0.01) and general health (p=0.02) in test 1 compared with responders in test 2 (table 1 and figure 1). Of the 554 participants, 166 completed web–web, 112 paper–paper, 239 web–paper and 37 paper–web, and the response rates in test 2 varied substantially between the four groups (figure 1). The median response time from test 1 to test 2 was 22 days (IQR 18 to 28 days).

Test–retest reliability and agreement of the PRO algorithm used as decision aid

Table 2 presents the agreement of the PRO algorithm used to identify patients with a need for contact with the outpatient clinic. Perfect algorithm agreement was observed in 82% of the cases (n=454). Disagreement was observed in 18%: 7% of the algorithms (n=39) changed status from yellow/red to green or red to yellow and 11% (n=61) changed status from green to yellow/red or yellow to red. Test–retest reliability and agreement estimates of the pooled PRO algorithm and in the different methods of administration are shown in table 3. Test–retest reliability in terms of the kappa statistic was borderline 'substantially' or 'moderate' in all methods of administration; however, there was a tendency towards higher estimates in similar method of administration (web–web or paper–paper) compared with mixed method of administration (web–paper or paper–web). Although the values varied, there was overlapping CIs among the groups (figure 2).

Test–retest reliability and agreement of single items

The test–retest reliability parameters of the single items included in the epilepsy questionnaire were moderate to substantial (online supplementary table 1). Test–retest reliability was fair to substantial in item categories within the framework of the PRO algorithm and perfect agreement ranged from 81.4% to 99.8%. Missing responses were less than 5% in all items. There was a skewed distribution in the majority of the item response scales, with high proportions of more than 15% at the upper or lower ends of the scale.

DISCUSSION

The PRO algorithm used to decide whether epilepsy outpatients need contact or not with the outpatient clinic has demonstrated substantial test–retest reliability: kappa with squared weight was 0.67 (95% CI 0.60 to 0.74). Perfect agreement was observed in 82% of the cases. There was a tendency towards higher test–retest reliability and agreement estimates within the same methods of administration (web–web or paper–paper) compared with a mixture of methods (web–paper or paper–web). For the majority of the included single items, kappa values were moderate

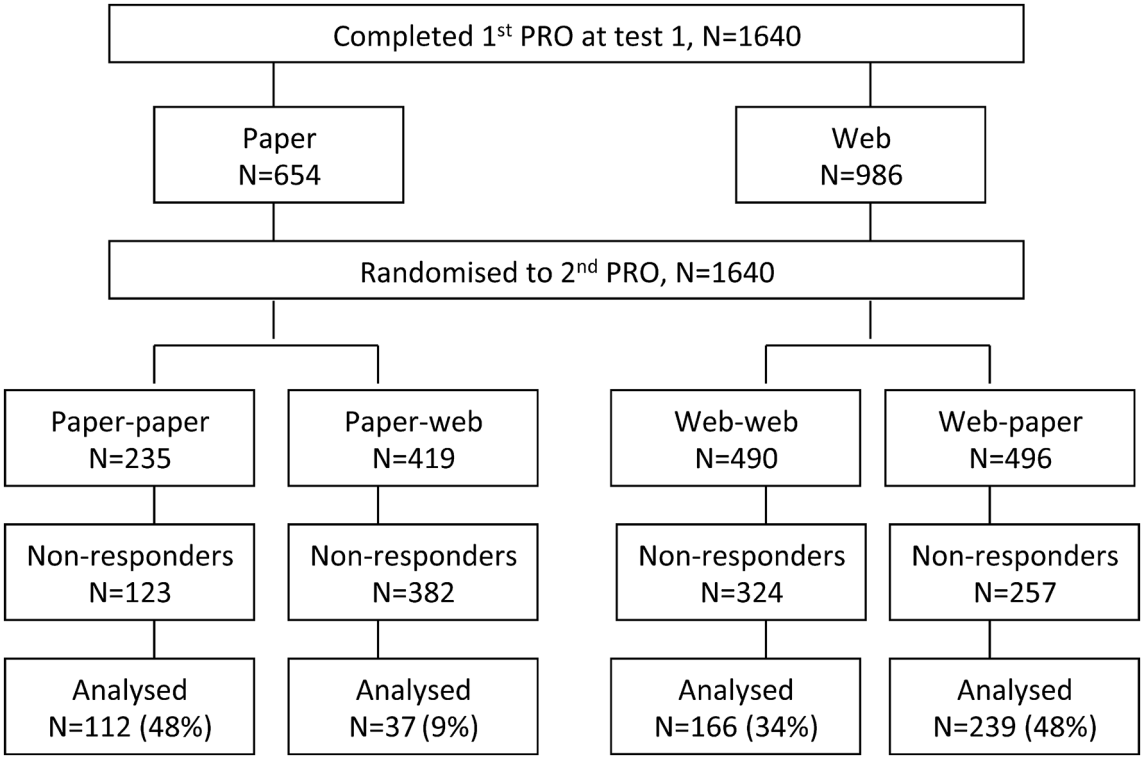


Figure 1 Flow chart of eligible participants' response method in test 1, randomisation of response method in test 2, non-responders in test 2 and participants included in the analysis. In paper responders, the randomisation allocation was 1:1 from August to November 2016, and 0.25:0.75 in favour of the web method from the end of November 2016 to April 2017. PRO, patient-reported outcome.

to substantial. Agreement exceeded 90%, whereas kappa values were fair to substantial in items within the framework of the PRO algorithm.

There are several sources of potential errors related to the consistency of a PRO measurement: (1) a real change in the patient health status between the two time points of measures, (2) difficulty related to answering items due to poor face validity and (3) incorrect answer from the patient made by mistake. Finally, the interval between the two measurement time points is important. A short interval increases the risk of recall bias and a long interval increases the risk of a real change in patient status.[24]

Table 2 Agreement between the automated PRO algorithm from test 1 to test 2, n=554

PRO algorithm test 1	PRO algorithm test 2			
	Green (%)	Yellow (%)	Red (%)	Total (%)
Green	104 (19)	42 (8)	1 (0.1)	147 (27)
Yellow	34 (6)	328 (59)	18 (3)	380 (69)
Red	0 (0)	5 (1)	22 (4)	27 (5)
Total	138 (25)	375 (68)	41 (7)	554 (100)

Green, no need of contact with the outpatient clinic.
Yellow, may need contact with the clinic (a clinician has to assess the PRO response).
Red, need of contact with the clinic.
PRO, patient-reported outcome.

This study found the highest test–retest reliability and agreement estimates in the web–web method of administration, however; not statistically significant from the paper–paper method. This finding is consistent with other studies which have reported that PRO data collected via the web method had the same quality as the paper-based method,[20 34 35] and in line with the recommendations from International Society for Pharmacoeconomics and Outcome Research (ISPORs) regarding electronic patient-reported outcome (ePRO); a web version of a paper version ought to produce data that are equivalent or superior.[36] Using the web-based method of PRO data collection has several advantages for patients as well as clinicians who use PRO data in clinical practice.[37] Egger *et al* evaluated the test–retest reliability of the Epidemiology of Prolapse and Incontinence Questionnaire in similar as well as mixed methods of administration and found no differences between the methods.[20] However, the tendency towards higher reliability and agreement estimates in similar method of administration compared with the mixed methods found in our study should be noted.

This study found a higher percentage of agreement in the worsening status of the PRO algorithm, indicating that the study population may have been less healthy in the second test of administration method. This finding was the same regardless of the methods of administration. This could have been caused by a real change in the participants' health status from test 1 to test 2. The

Table 3 Test–retest reliability and agreement between the PRO algorithm from test 1 to test 2 in the study population and in different methods of administration

PRO algorithm	n	Perfect agreement % (95% CI)	Disagreement improved status % (95% CI)	Disagreement worsening status % (95% CI)	Kappa* (95% CI)
Pooled	554	82 (78 to 85)	7 (5 to 9)	11 (9 to 14)	0.67 (0.60 to 0.74)
Web–web	166	87 (80 to 92)	5 (2 to 9)	8 (5 to 14)	0.78 (0.67 to 0.86)
Paper–paper	112	82 (74 to 89)	8 (4 to 15)	10 (5 to 17)	0.69 (0.57 to 0.81)
Mixed†	276	79 (74 to 84)	8 (5 to 12)	13 (9 to 18)	0.59 (0.48 to 0.69)

*Weighted Kappa with squared weights.
†Web–paper and paper–web.
PRO, patient-reported outcome.

interval period in this study was quite long in some participants, with a maximum range of 104 days and a median range of 22 days. This could potentially have caused bias if the disease status had changed. Therefore, subanalyses were made which tested whether the long interval had any impact on the overall estimates. The results showed a tendency towards an increase of the reliability estimates in similar method of administration, but a decrease in the mixed methods. Therefore, the difference may not be due to a real change in the participants' health status, but rather a consequence of the participants' response method. The participants self-selected the administration method in test 1, and a compulsory administration method in test 2 may be inconvenient and lead to biased answers. The different methods of administration and layout of the questionnaire in test 2 may have affected the participants' response habits, reflection or recall of the items, favouring identical methods.

Another limitation in this study was the risk of selection bias. The response rate was only 34%, ranging from 9% in the paper–web group to 48% in the paper–paper as well as the web–paper group. The low response rate was may caused by the pragmatic design where patients responded to their preferred method in test 1 as part of standard care in three outpatient clinics. The low response rate in the paper–web group compared with the paper–paper group could be related to the fact that the patient responders in test 1 had selected the paper method because of restricted access to respond via the internet. Furthermore, the use of reminders at test 2 could have increased the overall response rate; however, reminders were not used in this study due to the importance of the interval length between the two measurement points in a test–retest study. It would be preferable to randomise the response method in test 1 as well to make the groups more comparable. As shown in table 1 and figure 1, participants were more likely men, older and web responders. Furthermore, the participants had a tendency to have a less symptom burden, better general health and well-being, and less likely to have a red PRO algorithm compared with non-participants. This indicated that the study population was a healthier group of patients compared with the non-responders. A study population that does not represent the source population may entail problems with interpretation and generalisation of the results. In this study, the test–retest reliability may have been underestimated due to a healthy, stable and homogeneous study population.

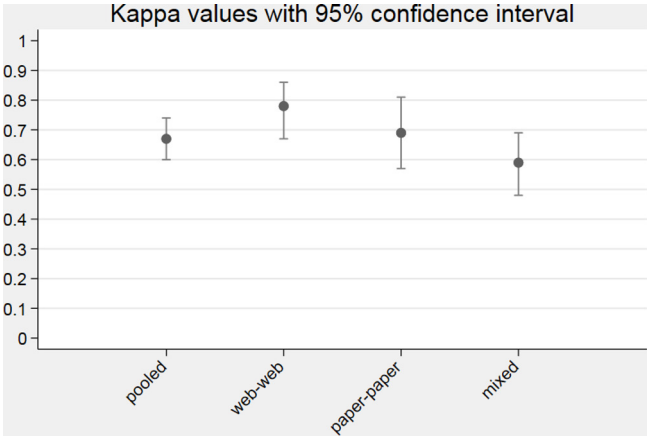


Figure 2 Test–retest reliability from test 1 to test 2 of the pooled PRO algorithm (n=554), web–web (n=166), paper–paper (n=112) and the mixed group (web–paper or paper–web, n=276). PRO, patient-reported outcome.

Kappa values are markedly affected by the prevalence of the measured event and distribution of item scores and a likely limitation of the interpretation of the results. This means that a high percentage agreement could potentially take place concurrent with a low kappa value if the prevalence of a specific item is low.[38] This was the case in the epilepsy questionnaire, in which a prevalence of less than 5% of the measured event was present in the majority of the items. For example, the two pregnancy items both had a low prevalence of the event. The percentage agreement was high, 99.6% and 98.9%, indicating a small measurement error; however, the kappa values were less convincing. Floor and ceiling effects could occur if a high proportion (more than 15%) of the respondents had a score at the lower or upper end of the scale.[33] This was the case in this study as well; concurrent with a low prevalence, a high proportion of the participants scored on the healthy side of the item response scales, indicating a homogeneous group.

This could potentially affect the reliability since it can be difficult to distinguish patients with the lowest or highest score from each other.[39] In addition, it could be difficult to measure longitudinal changes (responsiveness) in these patients as well.[39] These aspects should be taken into consideration in the interpretation of the kappa values.

CONCLUSION

This is the first test–retest reliability study of a disease-specific epilepsy PRO algorithm and questionnaire used to support clinical decision-making. In 2018, the questionnaire and the PRO algorithm are used by approximately 5000 patients with epilepsy in five outpatient clinics in Denmark. Overall, the PRO algorithm showed substantial test–retest reliability and agreements in same method of administration, whereas there was a tendency towards lower reliability and agreement if the method of administration was mixed.

Author affiliations
[1]AmbuFlex/West Chronic, Occupational Medicine, Regional Hospital West Jutland, University Research Clinic, Aarhus University, Herning, Denmark
[2]Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark
[3]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark
[4]Department of Occupational Medicine, Regional Hospital West Jutland, University Research Clinic, Herning, Denmark
[5]Department of Neurology, Aarhus University Hospital, Aarhus, Denmark
[6]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

Contributors NHH and LMVS designed the study protocol in collaboration with PS, AdT and DHC. LMVS participated in recruitment of participants, data collection, performed the statistical analyses and drafted the manuscript. NHH, AdT and DHC contributed to interpretation of data and critical revision of the manuscript. NHH, AdT, DHC and LMVS read and approved the final manuscript and stand by the integrity of the entire work.

Funding The study was funded by Aarhus University, the Central Denmark Regions Health Research Foundation and TrygFonden.

Competing interests None declared.

Patient consent Not required.

Ethics approval The study was approved by the Danish Data Protection Agency (j.no: 1-16-02-691-14). According to Danish law, approval by the ethics committee and written informed consent was not required. The eligible patients were provided with information about the study and its purpose, including that participation was voluntary.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement An anonymous version of the dataset used in this current study is available. Interested researchers may contact the corresponding author of this article for further guidance.

REFERENCES

1. US Department of Health and Human Services Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. 2009 www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf (assessed 2 Nov 2017).
2. Valderas JM, Kotzeva A, Espallargues M, *et al*. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 2008;17:179–93.
3. Santana MJ, Feeny D. Framework to assess the effects of using patient-reported outcome measures in chronic care management. *Qual Life Res* 2014;23:1505–13.
4. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract* 2006;12:559–68.
5. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res* 2013;13:211,6963–3.
6. Lohr KN, Zebrack BJ. Using patient-reported outcomes in clinical practice: challenges and opportunities. *Qual Life Res* 2009;18:99–107.
7. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014;23:508–18.
8. Wright JG. Evaluating the outcome of treatment. Shouldn't We be asking patients if they are better? *J Clin Epidemiol* 2000;53:549–53.
9. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, *et al*. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. *J Comp Eff Res* 2016;5:507–19.
10. Hjollund NH, Larsen LP, Biering K, *et al*. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. *Interact J Med Res* 2014;3:e5.
11. Snyder CF, Aaronson NK, Choucair AK, *et al*. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res* 2012;21:1305–14.
12. Christensen J, Vestergaard M, Pedersen MG, *et al*. Incidence and prevalence of epilepsy in Denmark. *Epilepsy Res* 2007;76:60–5.
13. Jennum P, Sabers A, Christensen J, *et al*. Welfare consequences for people with epilepsy and their partners: A matched nationwide study in Denmark. *Seizure* 2017;49:17–24.
14. Aguirre C, Quintas S, Ruiz-Tornero AM, *et al*. Do people with epilepsy have a different lifestyle? *Epilepsy Behav* 2017;74:27–32.
15. Fisher RS, van Emde Boas W, Blume W, *et al*. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia* 2005;46:470–2.
16. Moura LM, Schwamm E, Moura Junior V, Junior M V, *et al*. Feasibility of the collection of patient-reported outcomes in an ambulatory neurology clinic. *Neurology* 2016;87:2435–42.
17. Meyer AC, Dua T, Ma J, *et al*. Global disparities in the epilepsy treatment gap: a systematic review. *Bull World Health Organ* 2010;88:260–6.
18. Schougaard LM, Larsen LP, Jessen A, *et al*. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. *Qual Life Res* 2016;25:525–34.
19. Kottner J, Audigé L, Brorson S, *et al*. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96–106.
20. Egger MJ, Lukacz ES, Newhouse M, *et al*. Web versus paper-based completion of the epidemiology of prolapse and incontinence questionnaire. *Female Pelvic Med Reconstr Surg* 2013;19:17–22.
21. Sjöström M, Stenlund H, Johansson S, *et al*. Stress urinary incontinence and quality of life: a reliability study of a condition-specific instrument in paper and web-based versions. *Neurourol Urodyn* 2012;31:1242–6.
22. Nixon A, Kerr C, Breheny K, *et al*. Patient Reported Outcome (PRO) assessment in epilepsy: a review of epilepsy-specific PROs according to the Food and Drug Administration (FDA) regulatory requirements. *Health Qual Life Outcomes* 2013;11:38.
23. Patrick DL, Burke LB, Gwaltney CJ, *et al*. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. *Value Health* 2011;14:967–77.
24. Md FPM. *Quality of Life: The Assesment, Analysis and Interpretation of Patient-Reported Outcomes*. 2nd ed. West Sussex, England: John Wiley & Sons Ltd, 2007.
25. K: SJ. *The use of well-being measures in primary health care – the Dep-Care project; in World Health Organization, Regional Office for*

Europe: Well-Being Measures in Primary Health Care – the DepCare Project. Geneva: World Health Organization, 1998:E60246.
26. Bjorner JB, Thunedborg K, Kristensen TS, *et al*. The Danish SF-36 Health Survey: translation and preliminary validity studies. *J Clin Epidemiol* 1998;51:991–9.
27. Olsen LR, Mortensen EL, Bech P. The SCL-90 and SCL-90R versions validated by item response models in a Danish community sample. *Acta Psychiatr Scand* 2004;110:225–9.
28. Topp CW, Østergaard SD, Søndergaard S, *et al*. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom* 2015;84:167–76.
29. Bjorner JB, Damsgaard MT, Watt T, *et al*. Tests of data quality, scaling assumptions, and reliability of the Danish SF-36. *J Clin Epidemiol* 1998;51:1001–11.
30. Mokkink LB, Prinsen CA, Bouter LM, *et al*. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016;20:105–13.
31. Reichenheim ME. Confidence intervals for the kappa statistic. *The Stata Journal* 2004;4:421–8.
32. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
33. de Vet HC, Terwee CB, Mokkink LB. *Measurement in Medicine: a practical guide*. Cambridge: Cambridge University Press, 2011.
34. Bliven BD, Kaufman SE, Spertus JA. Electronic collection of health-related quality of life data: validity, time benefits, and patient preference. *Qual Life Res* 2001;10:15–21.
35. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health* 2008;11:322–33.
36. Coons SJ, Gwaltney CJ, Hays RD, *et al*. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009;12:419–29.
37. Jones JB, Snyder CF, Wu AW. Issues in the design of Internet-based systems for collecting patient-reported outcomes. *Qual Life Res* 2007;16:1407–17.
38. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360–3.
39. Terwee CB, Bot SD, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.

Appendix 1

The questionnaire and PRO-algorithm development process

A research consensus team that included clinical experts in epilepsy and experts in patient-reported outcome (PRO) measures was established in September 2011. The group provided inputs to the content and construct of the epilepsy questionnaire and the PRO-algorithm used to support clinical decision-making in outpatient follow-up.

PREPERATION

Rationale
Clinicians working with epilepsy experienced an increased volume of patients in the outpatient clinic and the majority of these patients were well treated. However, the need of monitoring treatment effect and screen for functional and mental health issues were still necessary. Therefore, self-reported data collected from the patients' home was assumed to have a great potential in this patient group.

Purpose
The purpose was to develop an instrument which could screen for epilepsy patients' health problems to support clinical decision-making in outpatient follow-up.

Patient group
The target group was patients with epilepsy ≥ 15 years with no cognitive impairments.

DEVELOPMENT

Conceptual framework of what do we want to be measured
Based on the Wilson and Cleary (1995) model, definition and levels of constructs were discussed. Symptom status, functional status, overall quality of life, and characteristics of the individual and the environment were identified as being clinically relevant aspects to measure in the population.

Development and selecting items
Literature
A systematic research strategy was conducted to identify establish generic and epilepsy-specific PRO instruments. Several instruments were identified; however, no established disease-specific instruments covering the purpose of identifying patients who need clinical attention were found. Several generic established instruments were identified covering some of the constructs that had been selected to be measured. These instruments included the World Health Organisation Well-Being Index (WHO-5), items from the Short Form 36 (SF-36), and items from the Symptom Checklist 92 (SCL-92).

Experts
Clinical experts in epilepsy, including both physicians and nurses, provided inputs to content development of new items. The severity of each symptom separately was considered important.

Only one aspect of the symptom was selected if the symptom has different aspects, e.g. fatigue. Experts in PRO provided inputs to formulating and scoring of the items.

TESTING AND EVALUATION

First draft: pilot testing
The first version of the questionnaire was pre-tested by using semi-structural interviewing techniques in 20 representative epilepsy patients from two outpatient clinics in Central Denmark Region. The aim of the pilot test was to identify potential problems such as low relevance of items, ambiguity of items, and lack of important topics.

Evaluation
The majority of the patients found the questionnaire content relevant, and no critical comprehension difficulties were identified. Some patients pointed out recall problems regarding some of the seizure items. They did not report lack of any essential topics nor did the time used to fill in the questionnaire raise any criticism.

THE PRO-ALGOITHM

Based on the first draft of the questionnaire, a clinical expert group in epilepsy assigned the response options for each item in three colours: green, yellow, or red. The allocation was based on what the doctors considered clinically important to react on to identify patients with need of attention. The overall aim of the algorithm is a 'red flag' approach. Specific answers result in a red colour indicating need of clinical attention, e.g. pregnancy, suicidal thoughts, seizure impairment.

A red colour indicates that the patient needs or wishes contact with the outpatient clinic. A yellow colour indicates that the patient may need contact with the clinic. In yellow cases, a clinician assesses the patient's PRO response, and based on the PRO data and other information in the patient's record it is decided whether further contact is needed. A green colour indicates that the patient does not need or wish contact with the clinic, and a subsequent questionnaire is sent to the patient at a pre-defined interval (e.g. after 3, 6, or12 months).

Subsequently, the PRO questionnaire and the PRO-algorithm were implemented and used in clinical practice, and experiences have been evaluated yearly since 2011 at consensus meetings.

Overview of items and the algorithm of the response categories in the epilepsy questionnaire

Item	**Response categories**					
Number of seizures during the last year	> 0					
Number of absence seizures during the last 3 months	> 0					
Number of generalised seizures during the last 3 months	> 0					
Seizure impairment	Yes	No				
Seizure injury	No	Yes, but not serious	Serious damage			
Emergency room visit due to epilepsy	Yes	No				
Relatives' worried	Never	Rarely	Occasionally	Frequently	Don't know	Not applicable
Headaches	Never	Occasionally	Sometimes	Often	Very often	
Dizziness	Never	Occasionally	Sometimes	Often	Very often	
Tremor/ shaking	Never	Occasionally	Sometimes	Often	Very often	
Double vision	Never	Occasionally	Sometimes	Often	Very often	
Loss of appetite	Never	Occasionally	Sometimes	Often	Very often	
Eating too much	Never	Occasionally	Sometimes	Often	Very often	
Difficulty remembering	Never	Occasionally	Sometimes	Often	Very often	
Difficulty concentrating	Never	Occasionally	Sometimes	Often	Very often	
Aggression	Never	Occasionally	Sometimes	Often	Very often	
Fatigue	Never	Occasionally	Sometimes	Often	Very often	
Sadness	Never	Occasionally	Sometimes	Often	Very often	
Fear of having seizures	Never	Occasionally	Sometimes	Often	Very often	
Problems with sexuality	Never	Occasionally	Sometimes	Often	Very often	
Being suicidal	Never	Occasionally	Sometimes	Often	Very often	
Well-being WHO-5 Index	Score < 50 or one extreme answer (at no time)					
General health	Excellent	Very good	Good	Fair	Poor	
General health compared to last year	Much better	Somewhat better	About the same	Somewhat worse	Much worse	
Medication adherence	Daily	Weekly	Monthly	Never / very rarely		
Side- effects	No	Yes, a few	Yes, some	Yes, many		
Work less because of epilepsy	Yes	Partly	No			
Social limitations	No	Yes				
Alcohol consumption	> consume 14/21 units a day					
Use of recreational drugs	Never	Monthly	Weekly	Daily		
Pregnant	Yes	No				
Planning pregnancy	Yes	No				
Car driving last month	Yes (+ seizures)	No				

Green: No need of contact with the outpatient clinic
Yellow: May need of contact (a clinician has to assess the PRO response)
Red: Need of contact with the clinic

Table 1 Agreement and reliability between the items from test 1 to test 2 in original categories and categories within the framework of the PRO- algorithm

		Original item categories			**Item categories within the framework of the PRO-algorithm**		
Item	*n/total*	Levels	Agreement/ Exp agreement %	KW^2 (95% CI)	Levels	Perfect/Exp agreement %	Unweighted kappa (95% CI)
Seizures last year	518/554	Continuous			2	94.2/62.4	0.85 (0.79; 0.90)
Absence seizures	516/554	Continuous			2	94.0/76.4	0.75 (0.66; 0.83)
Generalised seizures	525/554	Continuous			2	96.0/88.0	0.67 (0.53; 0.80)
Seizure impairment	154/135	2			2	90.3/79.8	0.52 (0.31; 0.73)
Seizure injury	156/135	3	94.1/85.4	0.59 (0.41; 0.79)	2	81.4/56.6	0.57 (0.43; 0.71)
ER visit	164/135	2			2	88.4/82.9	0.32 (0.09; 0.56)
Relatives' worried	517/554	6	92.8/85.9	0.49 (0.38; 0.59)	2	92.7/83.5	0.56 (0.43; 0.68)
Headaches	550/554	5	96.7/86.7	0.75 (0.69; 0.81)	2	92.2/82.2	0.56 (0.45; 0.68)
Dizziness	550/554	5	97.7/91.0	0.75 (0.69; 0.80)	2	96.6/91.8	0.58 (0.41; 0.75)
Tremor/ shaking	548/554	5	97.9/90.9	0.77 (0.70; 0.82)	2	96.4/91.0	0.60 (0.44; 0.76)
Double vision	550/554	5	97.8/93.4	0.66 (0.54; 0.75)	2	97.6/94.5	0.57 (0.36; 0.78)
Loss of appetite	551/554	5	98.2/94.4	0.68 (0.60; 0.77)	2	97.8/95.4	0.53 (0.29; 0.76)
Eating too much	546/554	5	97.0/89.8	0.71 (0.63; 0.77)	2	95.6/89.6	0.58 (0.42; 0.73)
Remembering	549/554	5	96.7/83.6	0.80 (0.75; 0.83)	2	91.1/73.5	0.66 (0.58; 0.75)
Concentrating	552/554	5	97.1/86.5	0.79 (0.74; 0.83)	2	95.3/83.2	0.72 (0.62; 0.82)
Aggression	551/554	5	97.0/90.1	0.69 (0.62; 0.75)	2	95.5/90.5	0.52 (0.36; 0.69)
Fatigue	549/554	5	96.0/83.3	0.76 (0.72; 0.80)	2	90.5/76.1	0.60 (0.51; 0.70)
Sadness	548/554	5	97.3/89.8	0.73 (0.68; 0.78)	2	95.6/89.0	0.60 (0.46; 0.75)
Fear of seizures	547/554	5	97.8/91.2	0.75 (0.69; 0.81)	2	96.3/91.3	0.58 (0.42; 0.75)
Sexuality	527/554	5	94.8/78.8	0.75 (0.69; 0.81)	2	86.0/62.5	0.63 (0.55; 0.70)
Suicidal	548/554	5	99.6/98.3	0.76 (0.60; 0.85)	3	99.1/97.0	0.69 (0.56; 0.83)
WHO-5 score	540/554	Continuous			2	90.4/75.8	0.60 (0.50; 0.70)
General health	552/554	5	97.1/88.7	0.74 (0.70; 0.78)	2	90.9/72.7	0.67 (0.58; 0.75)

General health last year	550/554	5	97.3/92.8	0.62 (0.54; 0.70)	2	91.5/81.9	0.53 (0.41; 0.65)
Medication adherence	545/554	4	97.3/92.9	0.62 (0.48; 0.74)	2	95.8/91.4	0.51 (0.33; 0.68)
Side-effects	538/554	4	96.5/88.1	0.71 (0.64; 0.77)	2	92.8/82.0	0.60 (0.48; 0.71)
Work less	240/164	3	94.0/87.8	0.51 (0.29; 0.69)	2	90.0/80.7	0.48 (0.31; 0.66)
Social limitations	518/554	2			2	91.1/75.8	0.63 (0.54; 0.73)
Alcohol female	254/268	Continuous			2	99.6/98.1	0.80 (0.41; 1.00)
Alcohol male	269/286	Continuous			2	98.1/94.6	0.66 (0.38; 0.94)
Recreational drugs	545/554	4	99.9/98.9	0.92 (0.57; 1.00)	2	99.8/99.1	0.80 (0.41; 1.00)
Pregnant	279/268	2			2	99.6/98.9	0.67 (0.05; 1.00)
Planning pregnancy	275/268	2			2	98.9/92.0	0.86 (0.71; 1.00)
Car driving	544/554	2			2	97.2/54.2	0.94 (0.91; 0.97)

Abbreviations: PRO, Patient-reported outcome; Exp, Expected; KW^2, Weighted kappa with squared weights; CI, Confidence interval; ER, Emergency room, WHO-5, WHO-5 Well-being Index

PAPER III

Schougaard et al. Health and Quality of Life Outcomes (2018) 16:175
https://doi.org/10.1186/s12955-018-1001-0

Health and Quality of Life Outcomes

SHORT REPORT Open Access

Test-retest reliability and measurement error of the Danish WHO-5 Well-being Index in outpatients with epilepsy

Liv Marit Valen Schougaard[1*], Annette de Thurah[2,3], Per Bech[4^], Niels Henrik Hjollund[1,5] and David Høyrup Christiansen[3,6]

Abstract

Background: The generic questionnaire WHO-5 Well-being Index (WHO-5), which measures the construct of mental well-being has been widely used in several populations across countries. The questionnaire has demonstrated sufficient psychometric properties; however, the test- retest reliability of the WHO-5 scale has yet to be determined. The aim of this study was to evaluate the test-retest reliability and measurement error of the Danish WHO-5 Well-being Index for outpatients with epilepsy. A further aim was to evaluate whether the method of administration (web, paper, or a mixture of the two modalities) influenced the results.

Methods: Epilepsy outpatients aged ≥15 years from three outpatient clinics in Central Denmark Region were included from August 2016 to April 2017. The participants were randomly divided into four test-retest groups: web-web, paper-paper, web-paper, and paper-web. Test-retest reliability was assessed by intraclass correlation coefficients (ICC) and measurement error by calculating minimal detectable change (MDC) on the basis of the standard error of the measurement.

Results: A total of 554 patients completed the questionnaire at two time points. The median duration between test-retest was 22 days. The pooled test-retest reliability estimate was ICC 0.81 (95% CI 0.78; 0.84). The estimated MDC was 23.60 points (95% CI 22.27; 25.10). These estimates showed little variation across administration methods.

Conclusions: WHO-5 showed acceptable test-retest reliability in a Danish epilepsy outpatient population across different method of administration; however, the relatively large measurement error should be taken into account when evaluating changes in WHO-5 scores over time. Further research should be done to explore these findings.

Keywords: Patient reported outcome measures, Validation studies as topic, Reproducibility of results

Introduction

Several considerations are important when selecting patient-reported outcome (PRO) measures for use in clinical practice. A PRO measure should be relevant to patients and clinicians and possess an adequate level of psychometric evidence for the instrument in the target population [1]. In Central Denmark Region, PRO measures have been used as the basis for follow-up in three epilepsy outpatient clinics since 2012 [2, 3]. Patients complete a web or paper-based questionnaire at home instead of having pre-scheduled appointments. Clinical resources could then be directed towards patients with actual need, and clinicians could use patients' self-reported information to identify otherwise undetected problems. As depression is common in patients with epilepsy [4], valid and reliable measurement tools are necessary to identify relevant symptoms. For this purpose, the WHO-5 Well-being Index (WHO-5) was selected and has been used since 2012 for outpatients with epilepsy in Central Denmark Region.

WHO-5 is a generic unidimensional questionnaire reflecting the construct mental well-being during the last 2 weeks [5]. The scale was developed in 1998 and has

* Correspondence: livschou@rm.dk
^Deceased
[1]AmbuFlex/WestChronic, Occupational Medicine, Regional Hospital West Jutland, University Research Clinic, Aarhus University, Herning, Denmark
Full list of author information is available at the end of the article

been widely used [6]. WHO-5 includes five positive wording statements rated on a 6-point ordinal scale ranging from 5 "all of the time" to 0 "at no time". Raw scores, which range from 0 to 25, are multiplied by 4 to obtain a percentage score ranging from 0 (worst) to 100 (best). A percentage score below 50 indicates poor mental well-being and a risk of depression. The WHO-5 has demonstrated sufficient psychometric properties in terms of construct validity, predictive validity, and internal consistency reliability in several patient populations including epilepsy [6–14]; however, the test- retest reliability of the WHO-5 scale has yet to be determined. Furthermore, few studies have explored the impact on consistency of using different methods of administration [15, 16].

The study aim was to evaluate the test-retest reliability and measurement error of the Danish WHO-5 Well-being Index for outpatients with epilepsy. A further aim was to evaluate whether the method of administration (web, paper, or a mixture of the two modalities) influenced the results.

Methods

Study population and setting

Patients with epilepsy aged ≥15 years from three outpatient clinics in Central Denmark Region were included from August 2016 to April 2017. The patients completed the questionnaire at two time points. First, they completed a questionnaire from the outpatient clinic based on their preferred web or paper administration method (test 1). Subsequently, approximately 2 weeks later, a letter was sent to the patients asking them to complete the same questionnaire again (test 2). The patients were randomly divided into four test-retest groups based on the method of administration at test 1 and test 2: web-web, paper-paper, web-paper, and paper-web. Three reminders were sent in test 1, but no reminders were sent to non-responders in test 2. The WHO-5 Well-being Index was included in the questionnaire in test 1. In addition, the questionnaire included other items, regarding, for example, seizures, symptoms, and general health. The general health construct was measured by using two items from the Danish version of the Short Form 36 Health Survey [17, 18]. A long interval between test administrations increases the risk of change in patients' health status in a test-retest study, whereas a short interval increases the risk of recall bias [19]. The questionnaire in test 1 was sent to the patients as part of routine outpatient follow-up. Patients' mental health was assumed to be stable during the time period from test 1 to test 2, since the health status of epilepsy patients is not likely to change over a period of 2 weeks. The patients were not asked in test 2 whether their mental health had changed within the time period.

Data analyses

Descriptive statistics were generated for patient characteristics and for each item to determine the extent of missing values and floor- or ceiling effects, which were considered present if more than 15% had a score at the lower or upper end of the scale [19]. Cronbach's alpha was used to assess internal consistency. The 95% confidence interval (CI) of the Cronbach's alpha values was estimated by using the bootstrap method (1000 replications). The time interval between test 1 and 2 was calculated as the difference in number of days from the dates of responses. Test-retest reliability of the scale was assessed by intraclass correlation coefficients (ICC) agreement model 2.1 [20], with 95% CI, and for single items, kappa with squared weights and 95% CI was used. An ICC value of 0.70 is considered acceptable for group level analysis, but when evaluating individual patients, an ICC of 0.90 is recommended [19]. The kappa values were interpreted as following: < 0.2 (slight), 0.21–0.4 (fair), 0.41–0.6 (moderate), 0.61–0.8 (substantial), and 0.81–1.0 (almost perfect) [21]. Measurement error was assessed with differences between test 1 and 2 plotted against the means of the two measurements by Bland–Altman plots with 95% CI and 95% limits of agreement (LOA). LOA equals the mean change in scores between test 1 and 2 (mean change ±1.96 x

Table 1 Patient characteristics at time of test 1 among outpatients with epilepsy, $N = 554$

Gender, *n* (%) Male	286	(52)
Age, *y*, median (IQR)	57.3	(25.1)
Outpatient clinic, *n* (%)		
Aarhus	409	(74)
Holstebro	115	(21)
Viborg	30	(5)
General health[a], *n* (%)		
Excellent	67	(12.1)
Very good	191	(34.5)
Good	209	(37.7)
Fair	68	(12.3)
Poor	19	(3.4)
WHO-5 score in test 1		
Median (IQR)	76	(24)
Mean (SD)	70.6	(19.5)
Missing, *n* (%)	5	(0.9)
WHO-5 score in test 2		
Median (IQR)	76	(24)
Mean (SD)	70.5	(19.2)
Missing, *n* (%)	9	(1.6)

[a]Item GH-1 from Short Form 36 Health Survey [17]
Abbreviations IQR inter quartile range, *SD* Standard deviation

Table 2 Item level distribution and weighted kappa of the WHO-5 Well-being Index ($N = 554$)

Item	Distribution (%) of the response options[a]								Test-retest Weighted kappa
	Item content		Missing	0	1	2	3	4	5	
1	I have felt cheerful and in good spirits	Test 1	0.2	0.5	5.2	5.6	12.3	61.6	14.6	0.70 (0.64; 0.76)
		Test 2	0.7	0.5	3.4	6.3	14.1	61.4	13.5	
2	I have felt calm and relaxed	Test 1	0.5	1.4	4.9	6.0	15.2	52.0	20.0	0.67 (0.59; 0.74)
		Test 2	0.5	1.3	2.7	7.9	13.4	56.9	17.3	
3	I have felt active and vigorous	Test 1	0.2	3.1	8.5	13.0	21.7	37.7	15.9	0.70 (0.65; 0.76)
		Test 2	0.5	2.9	9.0	13.5	19.7	41.9	12.5	
4	I woke up feeling fresh and rested	Test 1	0.2	4.7	10.6	10.5	18.8	40.4	14.8	0.72 (0.66; 0.77)
		Test 2	1.1	5.4	9.0	11.4	17.0	41.9	14.3	
5	My daily life has been filled with things that interest me	Test 1	0.5	0.9	5.4	7.0	16.6	51.3	18.2	0.68 (0.62; 0.74)
		Test 2	1.1	0.7	7.6	6.5	15.0	51.3	17.9	

[a]0 = At no time, 1 = Some of the time, 2 = Less than half of the time, 3 = More than half of the time, 4 = Most of the time, 5 = All of the time

standard deviation of the changes) and gives an indication of how much two scores can vary in stable patients. LOA are expressed in the units of measurement instrument and give a direct indication of the size of the measurement error [19]. The measurement errors reflect the within intraindividual variation and were estimated as the standard error of the measurement (SEM) [22]. SEM equals the square root of the error variance. The interpretation of a SEM estimate is not straight forward; therefore the SEM was converted into the minimally detectable change (MDC). MDC^{95} equals $1.96 \pm \sqrt{2}$ x SEM and indicates the smallest within-person change that can be interpreted as a "real" individual change above the measurement error [22]. Thus, a change in scores within the LOA or smaller than MDC^{95}can be attributed to measurement error [19]. Patients with missing item values were excluded from the analyses. Two sensitivity analyses were performed to investigate whether the length of the time interval between test 1 and test 2 affected our results. In the first analysis, patients were excluded if the time period between test 1 and test 2 was above 30 days, and in the second analysis all patients with a time interval above 14 days were excluded. STATA 15 software (Stata Corp, College Station) were used for all statistical analyses.

Results

Patient and item characteristics

A total of 554/1640 (34%) patients responded to the questionnaire twice. The median age was 57.3 years (Table 1). The response-rates in the four test-retest groups ranged from 48% (web-paper and paper-paper) to 34% (web-web) to 9% (paper-web). Non-responders were more likely younger, paper-responders, and had lower self-reported general health in test 1 (data not shown). The median response time between test-retest was 22 days (inter quartile range 10 days). A total of 14 patients had missing values for WHO-5 in test 1 or 2 and were excluded from the analyses. Percentages of missing values ranged from 0.2 to 1.1%, and there was a tendency towards ceiling effects in all items (Table 2). Cronbach's alpha was 0.89 (95% CI 0.87; 0.90) in test 1 and 0.89 (95% CI 0.87; 0.91) in test 2.

Test-retest reliability and measurement error of WHO-5

Kappa values for the five single items were substantial (Table 2) [21]. The ICC of the pooled WHO-5 score was 0.81 (95% CI 0.78; 0.84) (Table 3). Differences between test 1 and test 2 plotted against the mean of the two tests with upper and lower LOAs are shown in Fig. 1. The estimated SEM was 8.51 points (95% CI 8.03; 9.05),

Table 3 Test-retest reliability and measurement error for the WHO-5 Well-being Index between test 1 and test 2

WHO-5	N	Mean, (95% CI) Test 1	Mean (95% CI) Test 2	Difference (95% CI)	SEM (95% CI)	ICC (95% CI)	MDC^{95} (95% CI)
Pooled	540	70.58 (68.94; 72.21)	70.40 (68.78; 72.02)	0.18 (−0.84; 1.20)	8.51 (8.03; 9.05)	0.81 (0.78; 0.84)	23.60 (22.27; 25.10)
Web-web	164	69.83 (66.73; 72.93)	70.10 (67.01; 73.18)	−0.27 (−2.02; 1.49)	8.05 (7.26; 9.03)	0.84 (0.80; 0.89)	22.31 (20.13; 25.03)
Paper-paper	107	70.65 (66.41; 74.90)	70.69 (66.87; 74.51)	−0.04 (− 2.56; 2.49)	9.31 (8.21; 10.76)	0.81 (0.74; 0.87)	25.81 (22.76; 29.82)
Web-paper	233	71.10 (68.85; 73.33)	70.73 (68.37; 73.09)	0.36 (−1.17; 1.89)	8.36 (7.66; 9.20)	0.78 (0.73; 0.83)	23.18 (21.24; 25.49)
Paper-web	36	70.44 (63.63; 77.26)	68.78 (62.05; 75.50)	1.67 (−2.79; 6.12)	9.30 (7.55; 12.14)	0.78 (0.66; 0.91)	25.79 (20.92; 33.64)

Abbreviations: *WHO-5* WHO-5 Well-being Index, *N* Number, *CI* Confidence Interval, *SEM* Standard error of the measurement, *ICC* Intra class correlation coefficient, *MDC* Minimal detectable change

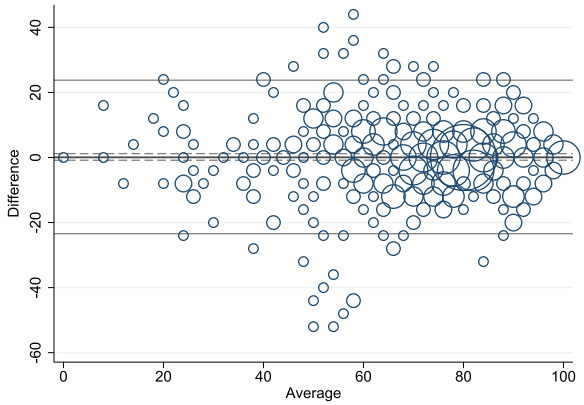


Fig. 1 Differences in the WHO-5 Well-being Index score between test 1 and test 2 plotted against the mean, $N = 540$

which resulted in a MDC^{95} of 23.60 points (95% CI 22.27; 25.10). The analysis was repeated in the four test-retest groups (Table 3 and Fig. 2). Administration methods did not noticeably alter the estimates. The overall results did not change, when the analyses were repeated with restricted intervals between test 1 and 2.

Discussion

Test-retest reliability of the Danish WHO-5 Well-being Index was found to be acceptable in an epilepsy outpatient population, but a relatively large measurement error was observed. The estimated MDC^{95} was 23.60 points, indicating that changes in the WHO-5 instrument must be substantial to ensure that a 'real' change is not due to measurement error. Methods of administration did not markedly influence the results.

This study follows the COSMIN framework [23, 24] and supplements earlier established psychometric properties of the WHO-5. Since we were unable to identify other test-retest studies of the scale, we believe this is the first study to determine the test-retest reliability of the WHO-5. Several studies have explored another aspect of reliability: internal consistency [8–14]. The Cronbach's alpha of the WHO-5 in these studies ranged from 0.82 to 0.95, which is consistent with the findings in this study. However, this aspect determines the correlation between items within a scale and not the degree of agreement for repeated measurements over time [22, 24]. The unidimensionality of the WHO-5 scale has been confirmed by using Rasch item response theory analyses in both a younger and elderly population [14, 25].

Test-retest reliability should be assessed in a stable population with an appropriate time interval between measurements [22]. We assumed that the epilepsy outpatient population was stable and allowed a longer time interval. Sensitivity analyses were used to assess potential change in health status; however, excluding participants with longer intervals between test 1 and 2 did not substantially alter the estimates. Still, we cannot rule out that a change in patients' health status had occurred and that this might have affected the ICC and measurement error estimates of the WHO-5 scale, as we did not collect information on the change in patients' mental health status from test 1 to test 2.

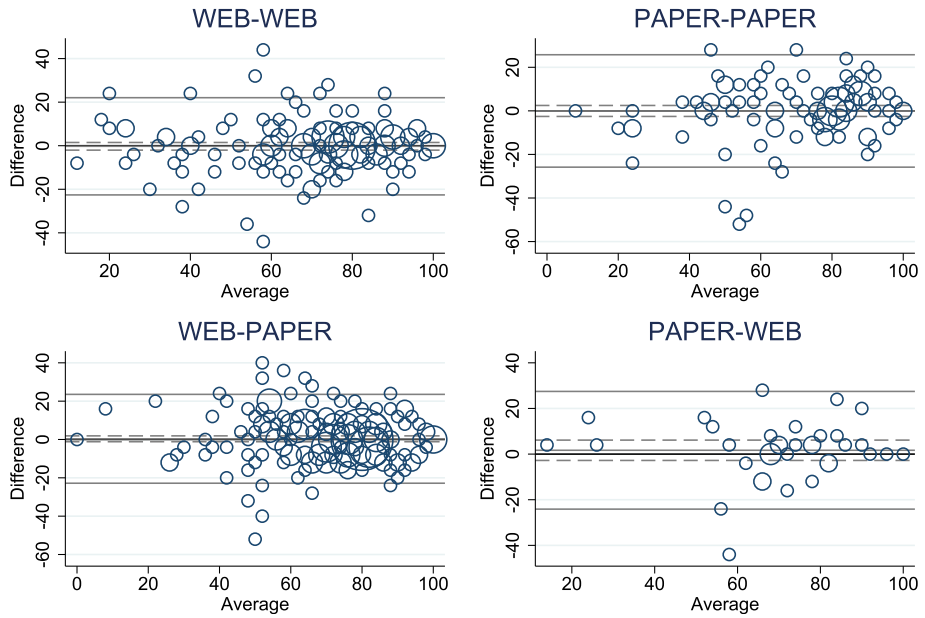


Fig. 2 Differences in the WHO-5 Well-being Index score between test 1 and test 2 plotted against the mean in the four test-retest groups: web-web ($n = 164$), paper-paper ($n = 107$), web-paper ($n = 233$), and paper-web ($n = 36$)

The WHO-5 scale ranges from 0 to 100, and an MDC of 23.6 points observed in this study may indicate that longitudinal differences of at least 24 points are needed to detect a "true" within-person change. The relatively large measurement error observed in this study may be taken into consideration by researchers planning future clinical trials and clinicians who use the scale on the individual level in clinical practice to evaluate change over time. Furthermore, the tendency towards ceiling effect may produce difficulties in measuring longitudinal changes. Web, paper, or a mixture of the two modalities showed nearly the same test-retest reliability, which is consistent with other test-retest studies [15, 16].

One important limitation of this study is the possibility of selection bias. A very low response rate was observed especially in the paper-web group (9%). This may be due to the pragmatic design, which allowed patients to choose administration method for their response to test 1. In the Danish general population, a mean WHO-5 score of 70 points has been reported [26, 27]. This is comparable with the result in this study; however, the responders tended to be a healthier group of patients compared to non-responders in test 2 who had lower self-reported general health and mental well-being in test 1. The reliability estimates indicate how well patients can be distinguished from each other despite the presence of measurement error, e.g. a lower ICC value tends to occur in a homogenous study sample [19]. Thus, in this study, the ICC estimates may have been underestimated due to a homogenous and healthy study population; whereas the measurement error estimates were probably less affected.

Conclusion

The WHO-5 Well-being Index showed acceptable test-retest reliability in a Danish epilepsy outpatient population, but the measurement error of the scale was relatively large. Different methods of administration did not influence the results. Further studies are required to provide insight into the test-retest reliability and measurement error in different language versions of the WHO-5 Well-being Index and in different patient populations.

Abbreviations

CI: Confidence Interval; ICC: Intraclass correlation coefficients; IQR: Inter quartile range; LOA: Limits of agreement; MDC: Minimal detectable change; PRO: Patient-reported outcome; SD: Standard deviation; SEM: Standard error of the measurement; WHO-5: WHO-5 Well-being Index

Funding

This study is funded by Aarhus University, the Health Research Fund of Central Denmark Region, and the Danish foundation TrygFonden. The funding sources had no role in the design of the study, data collection, analysis, interpretation of data, and writing the manuscript.

Availability of data and materials

An anonymous version of the datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

DHC, NHH, AdT and LMVS conceived the study. LMVS participated in data collection, performed the statistical analyses, and drafted the manuscript. All authors contributed to interpretation of data and critical revision of the manuscript. All authors read and approved the final manuscript and stand by the integrity of the entire work.

Ethics approval and consent to participate

The study was approved by the Danish Data Protection Agency (j.no: 1–16–02-691-14). All procedures performed were in accordance with the ethical standards of national research committee and with the 1964 Declaration of Helsinki. According to Danish law, approval by the ethics committee and written informed consent was not required. The eligible patients were provided with information about the study and its purpose, including that participation was voluntary.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]AmbuFlex/WestChronic, Occupational Medicine, Regional Hospital West Jutland, University Research Clinic, Aarhus University, Herning, Denmark. [2]Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark. [3]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark. [4]Psychiatric Research Unit, Psychiatric Centre North Zealand, University of Copenhagen, Hillerød, Denmark. [5]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark. [6]Department of Occupational Medicine, Regional Hospital West Jutland, University Research Clinic, Herning, Denmark.

Received: 19 February 2018 Accepted: 28 August 2018
Published online: 06 September 2018

References

1. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, Hess R, Miller DM, Reeve BB, Santana M. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Qual Life Res. 2012;21(8):1305–14.
2. Schougaard LM, Larsen LP, Jessen A, Sidenius P, Dorflinger L, de Thurah A, Hjollund NH. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Qual Life Res. 2016; 25(3):525–34.
3. Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjaer E, Schougaard LM. Use of patient-reported outcome (PRO) measures at group and patient levels: experiences from the generic integrated PRO system WestChronic. Interact J Med Res. 2014;3(1):e5.
4. Fiest KM, Dykeman J, Patten SB, Wiebe S, Kaplan GG, Maxwell CJ, Bulloch AG, Jette N. Depression in epilepsy: a systematic review and meta-analysis. Neurology. 2013;80(6):590–9.
5. WHO Regional Office for Europe: Wellbeing measures in primary health care / the DepCare project; in World Health Organization. World Health Organization. 1998. http://www.euro.who.int/__data/assets/pdf_file/0016/130750/E60246.pdf. Assessed 18 Dec 2017.
6. Topp CW, Ostergaard SD, Sondergaard S, Bech P. The WHO-5 well-being index: a systematic review of the literature. Psychother Psychosom. 2015; 84(3):167–76.
7. Hansen CP, Amiri M. Combined detection of depression and anxiety in epilepsy patients using the neurological disorders depression inventory for epilepsy and the World Health Organization well-being index. Seizure. 2015; 33:41–5.
8. Halliday JA, Hendrieckx C, Busija L, Browne JL, Nefs G, Pouwer F, Speight J. Validation of the WHO-5 as a first-step screening instrument for depression in adults with diabetes: results from diabetes MILES - Australia. Diabetes Res Clin Pract. 2017;132:27–35.

9. Newnham EA, Hooke GR, Page AC. Monitoring treatment response and outcomes using the World Health Organization's wellbeing index in psychiatric care. J Affect Disord. 2010;122(1–2):133–8.
10. Guethmundsdottir HB, Olason DP, Guethmundsdottir DG, Sigurethsson JF. A psychometric evaluation of the Icelandic version of the WHO-5. Scand J Psychol. 2014;55(6):567–72.
11. Krieger T, Zimmermann J, Huffziger S, Ubl B, Diener C, Kuehner C, Grosse Holtforth M. Measuring depression with a well-being index: further evidence for the validity of the WHO well-being index (WHO-5) as a measure of the severity of depression. J Affect Disord. 2014;156:240–4.
12. Hajos TR, Pouwer F, Skovlund SE, Den Oudsten BL, Geelhoed-Duijvestijn PH, Tack CJ, Snoek FJ. Psychometric and screening properties of the WHO-5 well-being index in adult outpatients with type 1 or type 2 diabetes mellitus. Diabet Med. 2013;30(2):e63–9.
13. de Wit M, Pouwer F, Gemke RJ, Delemarre-van de Waal HA, Snoek FJ. Validation of the WHO-5 well-being index in adolescents with type 1 diabetes. Diabetes Care. 2007;30(8):2003–6.
14. Lucas-Carrasco R. Reliability and validity of the Spanish version of the World Health Organization-five well-being index in elderly. Psychiatry Clin Neurosci. 2012;66(6):508–13.
15. Egger MJ, Lukacz ES, Newhouse M, Wang J, Nygaard I. Web versus paper-based completion of the epidemiology of prolapse and incontinence questionnaire. Female Pelvic Med Reconstr Surg. 2013;19(1):17–22.
16. Sjostrom M, Stenlund H, Johansson S, Umefjord G, Samuelsson E. Stress urinary incontinence and quality of life: a reliability study of a condition-specific instrument in paper and web-based versions. Neurourol Urodyn. 2012;31(8):1242–6.
17. Bjorner JB, Thunedborg K, Kristensen TS, Modvig J, Bech P. The Danish SF-36 health survey: translation and preliminary validity studies. J Clin Epidemiol. 1998;51(11):991–9.
18. Bjorner JB, Damsgaard MT, Watt T, Groenvold M. Tests of data quality, scaling assumptions, and reliability of the Danish SF-36. J Clin Epidemiol. 1998;51(11):1001–11.
19. de Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: a practical guide. 1st ed. United Kingdom: Cambridge University Press; 2011.
20. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.
22. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
23. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN risk of Bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5): 1171–79.
24. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737–45.
25. Blom EH, Bech P, Hogberg G, Larsson JO, Serlachius E. Screening for depressed mood in an adolescent psychiatric context by brief self-assessment scales--testing psychometric validity of WHO-5 and BDI-6 indices by latent trait analyses. Health Qual Life Outcomes. 2012;10:149–7525. 10-149
26. Bech P, Olsen LR, Kjoller M, Rasmussen NK. Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 mental health subscale and the WHO-five well-being scale. Int J Methods Psychiatr Res. 2003;12(2):85–91.
27. Ellervik C, Kvetny J, Christensen KS, Vestergaard M, Bech P. Prevalence of depression, quality of life and antidepressant treatment in the Danish general suburban population study. Nord J Psychiatry. 2014;68(7):507–12.

PAPER IV

Including online supplemental materials:

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents.

Schougaard et al. BMC Health Services Research (2017) 17:83
DOI 10.1186/s12913-017-2015-8

BMC Health Services Research

STUDY PROTOCOL Open Access

Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: study protocol for a pragmatic randomised controlled study

Liv Marit Valen Schougaard[1*], Caroline Trillingsgaard Mejdahl[2,6], Klaus Hvam Petersen[1], Anne Jessen[1], Annette de Thurah[2,3,4], Per Sidenius[5], Kirsten Lomborg[2,4,6] and Niels Henrik Hjollund[1,7]

Abstract

Background: The traditional system of routine outpatient follow-up of chronic disease in secondary care may involve a waste of resources if patients are well. The use of patient-reported outcomes (PRO) could support more flexible, cost-saving follow-up activities. *AmbuFlex* is a PRO system used in outpatient follow-up in the Central Denmark Region. PRO questionnaires are sent to patients at fixed intervals. The clinicians use the PRO data to decide whether a patient needs a visit or not (standard telePRO). PRO may make patients become more involved in their own care pathway, which may improve their self-management. Better self-management may also be achieved by letting patients initiate contact. The aim of this study is to obtain data on the effects of patient-initiated follow-up (open access telePRO) on resource utilisation, quality of care, and the patient perspective.

Methods: The study is a pragmatic, randomised, controlled trial in outpatients with epilepsy. Participants are randomly assigned to one of two follow-up activities: a) standard telePRO or b) open access telePRO. Inclusion criteria are age ≥ 15 years and previous referral to standard telePRO follow-up at Aarhus University Hospital, Denmark. Furthermore, patients must have answered the last questionnaire via the Internet. The number of contacts will be used as the primary outcome measure. Secondary outcome measures include well-being (WHO-5 Well-Being Index), general health, number of seizures, treatment side effects, mortality, health literacy (Health Literacy Questionnaire), self-efficacy (General Self-Efficacy scale), patient activation, confidence, safety, and satisfaction. In addition, the patient perspective will be explored by qualitative methods. Data will be collected at baseline and 18 month after randomisation. Inclusion of patients in the study started in January 2016. Statistical analysis will be performed on an intention-to-treat and per-protocol basis. For qualitative data, the interpretive description strategy will be used.

Discussion: The benefits and possible drawbacks of the PRO-based open access approach will be evaluated. The present study will provide important knowledge to guide future telePRO interventions in relation to effect on resource utilisation, quality of care, and the patient perspective.

Trial registration: ClinicalTrials.gov: NCT02673580 (Registration date January 28, 2016)

Keywords: Patient-reported outcomes, TelePRO, Clinical practice, Outpatient clinic, Outpatient follow-up, Open access, Randomised controlled trial

* Correspondence: livschou@rm.dk
[1]AmbuFlex, Regional Hospital West Jutland, Herning, Denmark
Full list of author information is available at the end of the article

Background

The Danish health care system is changing from inpatient towards a greater outpatient activity. From 2002 to 2009, there was a 50% increase in outpatient activity in Denmark, primarily related to the number of contacts per patient [1]. At the same time, there appears to be a growing need of health care services especially for the growing group of patients with chronic diseases and an increased focus on patient involvement. The challenge is to manage this without compromise on quality of care and patient outcomes. Follow-up visits for patients with chronic diseases in secondary care are traditionally based on regular pre-booked visits, which may be arranged when the patient is well. Thus patients as well as clinicians may find such visits unnecessary. The volume of appointments leads to capacity issues in outpatient clinics that struggle to respond rapidly to patients' requests for help [2].

One way of handling this challenge may be to let patients report essential information on health status and symptoms from home before or instead of visiting the outpatient clinic. Patients' own reports on health condition are termed patient-reported outcomes (PRO). The American Food and Drug Agency definition of PRO, *"A measurement based on a report that comes directly from the patient about the status of a patient's health condition without interpretation of the patient's response by a clinician or anyone else"* [3], focuses on the source of information and points out the importance of the patient perspective. The use of PRO in clinical practice is becoming increasingly common, and studies have reported improved patient-clinician communication, more effective self-management, and better utilisation of resources when PROs are used, whereas findings related to effects on patient outcomes are less consistent [4–7].

PRO may facilitate patient involvement because the problems reported as important by the patient are taken into consideration in the decision-making process [8–10]. However, patient involvement is not a goal in itself but rather a means to increase the patient's self-management. Self-management refers to the individual's ability to manage symptoms, treatment, physical and psychosocial consequences, and life style changes inherent in life with a chronic disease [11]. In practice, PRO is supposed to promotes a patient-centred dialogue between the patient and the clinicians in which the patient's view and opinion on his health are included. Thus, implementing PRO into clinical practice allows patients to actively participate in their own care, by which their self-management may improve [8].

AmbuFlex is a generic clinical PRO system which is not limited to specific patient groups, organisations or medical record systems [12]. As of December 2015, AmbuFlex had been implemented in nine patient groups at 15 outpatient clinics in Denmark [13]. An analysis initiated by the Danish government based on experiences with AmbuFlex has demonstrated a positive national business case and considerable quality gain [14]. The Danish government and Danish regions, who run the public hospitals, have decided on an agreement for nationwide implementation of PRO in three diagnostic group, including epilepsy, before 2020. AmbuFlex was implemented for epilepsy outpatients at Aarhus University Hospital in March 2012 and is now used at three neurological departments in the Central Denmark Region. As of August 2016, 4,513 epilepsy outpatients have been referred to AmbuFlex, which are about two-thirds of all epilepsy outpatients in the region. The PRO questionnaire used contains information on specific aspects of daily life with epilepsy and has been developed in close cooperation with clinicians and patients. Face validity is fundamental and has been ensured during the development of the questionnaire [12, 13]. A graphical PRO overview is presented to the clinicians, who use the PRO data for clinical decisions together with other available clinical data in the record to decide whether the patient needs a visit or not. If a PRO questionnaire is used to evaluate the patient's need for a hospital visit, the PRO data must be obtained outside the hospital. This is called tele-patient-reported outcome (telePRO) [12]. Experiences from epilepsy outpatient clinics have shown that of 8,256 PRO-based contacts, 48% were handled without additional contact to the patient other than the PRO questionnaire [13]. A preliminary interview study has indicated that patients experience greater flexibility in care, the saving of time, improved communication with the clinicians, and increased knowledge about their own disease [13, 15].

The AmbuFlex method used at the three neurological departments is called *standard telePRO.* In standard telePRO, regular scheduled visits are replaced with fixed questionnaires at intervals similar to those of the former pre-booked visits. A patient-initiated approach "open access" telePRO has been developed in which patients have access to their own PRO data and are able to initiate contact with the clinic by filling in a PRO questionnaire. A review by Whear et al. investigated the effectiveness of patient-initiated clinics in chronic conditions in secondary care and included seven randomised trials. The review found that the risk of harm from using the patient-initiated clinic model is low in patients with breast cancer, inflammatory bowel disease, and rheumatoid arthritis. The included studies found few significant differences in clinical outcomes between traditional appointment scheduling and the patient-initiated follow-up method. In four of the studies, the patient-initiated model was associated with savings in clinician time and resource use [2]. A review by Taneja et al. that included five of the same randomised studies reached the same

conclusion [16], while another review showed no significant differences in psychological and health-related quality of life outcomes between consultant-led and patient-initiated clinics. Patients have reported better satisfaction in patient-initiated clinics compared to usual care [17]. The patient-initiated method used was broadly the same in the studies included in the three reviews. Patients could request clinical advice by calling the clinic and, if necessary, arranging an appointment to see a clinician. However, none of the included studies used PRO as the main access point in the open access intervention, and all studies contain methodological limitations [2, 16, 17].

Objectives

The aim of this study is to provide insight into the effects of patient-initiated telePRO follow-up. The specific aims are to compare resource utilisation, quality of care, and the patient perspective of two outpatient follow-up activities: a) standard telePRO (fixed-interval telePRO follow-up) and b) open access telePRO (patient-initiated telePRO follow-up). We hypothesise that 1. Number of contacts is less in open access telePRO, 2. Quality of care in open access telePRO is at least as good as in standard telePRO, and 3. Patient self-management and experiences in open access telePRO are better than standard telePRO.

Methods

The study follows the (Additional file 1: SPIRIT checklist): Standard protocol items for clinical trials [18].

Design

This study is a pragmatic two-arm randomised controlled trial. Participants are randomly assigned to one of two follow-up activities: (a) standard telePRO or (b) open access teleRPO.

Study population

Participants are epilepsy outpatients recruited from the epilepsy clinic at Aarhus University Hospital in Central Denmark Region, Denmark.

Inclusion criteria

a) Age ≥ 15 years
b) Diagnosis or suspicion of epilepsy (IC-D 10 codes: G40, Z033a, DR568 and DR568E)
c) Already referred to standard telePRO by a clinician
d) Able to answer the questionnaire via the Internet, indicated by having answered the last questionnaire via the Internet

Exclusion criterion

a) Referred to telePRO follow-up with proxy questionnaire. Patients can be referred to a proxy questionnaire if they have cognitive problems and need help from a relative or health professionals.

Intervention

Reference group – standard telePRO

AmbuFlex (standard telePRO) is used in three epilepsy outpatient clinics in Central Denmark Region. In standard telePRO, outpatient follow-up activity is determined by a clinician and patients receive a questionnaire at fixed intervals (3, 6, or 12 months). The questionnaire includes information about aspect of daily life with epilepsy such as seizures, symptoms, medication adherence, and social aspects. Responses are automatically processed according to a specific algorithm and given a "green", "yellow", or "red" status. A red status indicates that the patient needs or wishes contact with the clinic, a green status that the patient has no current need of attention, while a yellow status indicates that the patient may need to be seen in the clinic, but a clinician has to decide whether further contact is needed. The patient can always overrule a decision by requesting contact. They can choose two different contact forms in the questionnaire: telephone consultation or a face-to-face consultation at the clinic. Non-responders get three reminders and are contacted if do not respond. Clinicians keep track of incoming yellow and red responses, and non-responders, and this information is presented on a PRO alert list. The PRO overview (Fig. 1) is presented graphically to the clinician within the electronic health record system, and used as decision aid together with other available health record information to decide whether the patient needs a visit or not [13].

Intervention group – open access telePRO

In open access, contact to the outpatient clinic is initiated by the patient by filling in a PRO questionnaire. The same questionnaire is used as in standard telePRO, but the patient decides when to respond. The patients can access a PRO overview, "My Epilepsy", customised for patient use via a secure login at the Danish national health website "Sundhed.dk". The clinicians handle questionnaires in the same way as in standard telePRO.

The open access website "My Epilepsy": design and features

A prototype website, "My Epilepsy", was developed to collect PRO in patient-initiated outpatient follow-up. The website was linked with the Danish National Health Website 'Sundhed.dk'. The website, "My Epilepsy", was customised for patient use and designed to allow patients to: a) answer a PRO questionnaire to get in contact with the clinic, b) view their personal PRO data (previously questionnaire responses), c) view information about the epilepsy questionnaire and specific questions, and d) have access to contact information to

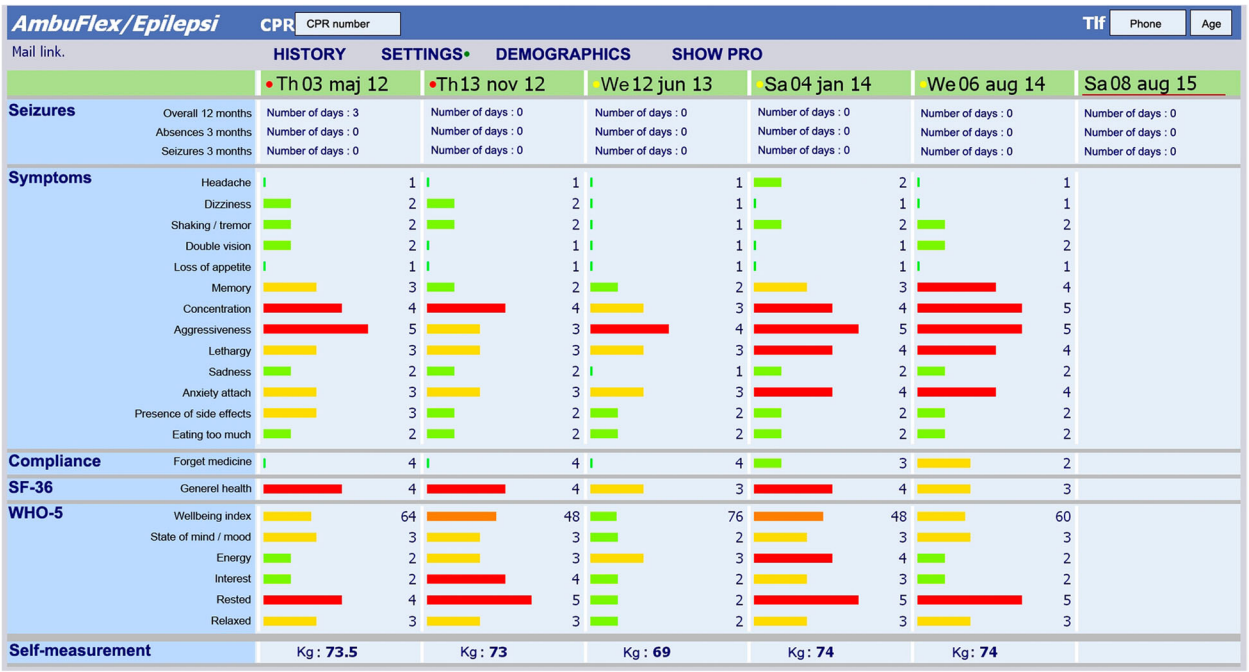


Fig. 1 Screen capture of the clinicians' overview in epilepsy clinics accessed from the Electronic Health Record of Central Denmark Region [13]. The colour dots in the upper row indicate the result of the automated PRO algorithm (red: definite need of contact, yellow: possible need of contact, green: no need of contact). Note that the colours of the bars have different meanings. The bars indicate the severity of the symptoms reported by the patient. A red or orange bar indicates a self-reported problem, a yellow bar some problem, and a green bar indicates no problems. Note: Labels were translated from Danish

the epilepsy outpatient clinic. A research team that included, outpatients with epilepsy and experts in telePRO, patient involvement, software technology, clinical epilepsy, provided inputs to design the prototype website. The research team developed the initial website specifications, constructed the website, and elicited feedback from epilepsy outpatients ($n = 6$), using cognitive interviewing techniques to study the manner in which the patients understood and responded to the website. The interface is shown in Fig. 2. Patients emphasised the importance of a user-friendly interface with clear and concise information. Patients were interested in tracking change over time and in using the website because it gave them the potential to communicate with their clinicians at a time decided by themselves. They found it conceivable that access to their previous questionnaires could give them a better understanding of their chronic disease. Finally, they pointed out the need for a telephone number if they required immediate contact. Patients had few problems assessing and using the site.

The website consists of four core elements:

a) Answer questionnaire: Here, patients can answer the epilepsy questionnaire when they need to get in contact with the clinic. The questionnaire is the same as in standard telePRO. When the patient has completed the questionnaire, the response is automatically sent as a red request to the PRO alert list at the epilepsy clinic. The clinician assesses the response and contacts the patient as soon as possible. The clinic has reserved appointments in their booking system to ensure that patients get a quick appointment. As a "safety net", patients have to answer the questionnaire before twice the fixed interval has elapsed. For example, if the patient is referred with a 12-month interval, the patient has to respond within two years. If not, the patient is automatically sent a questionnaire, given a red status, and is contacted by a clinician.

b) Previous answers: In this element, all of the patient's previous questionnaire responses are available. Patients have access to a PRO overview interface and specific and detailed questionnaire responses in the same manner as the clinicians. The overview interface is shown in Fig. 3. It is customised to monitor selected PRO data and to illustrate changes in health status over time. Colour codes indicate the severity of the symptoms reported by the patient. A red or orange bar indicates a self-reported problem, a yellow bar some problem, and green bar indicates no problems.

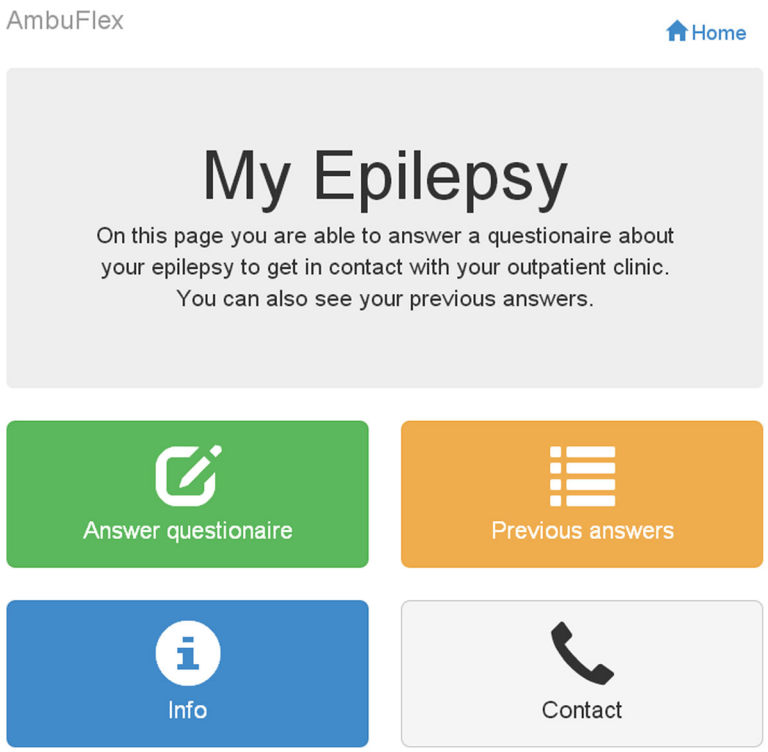


Fig. 2 The open access telePRO website "My Epilepsy". Note: Labels were translated from Danish

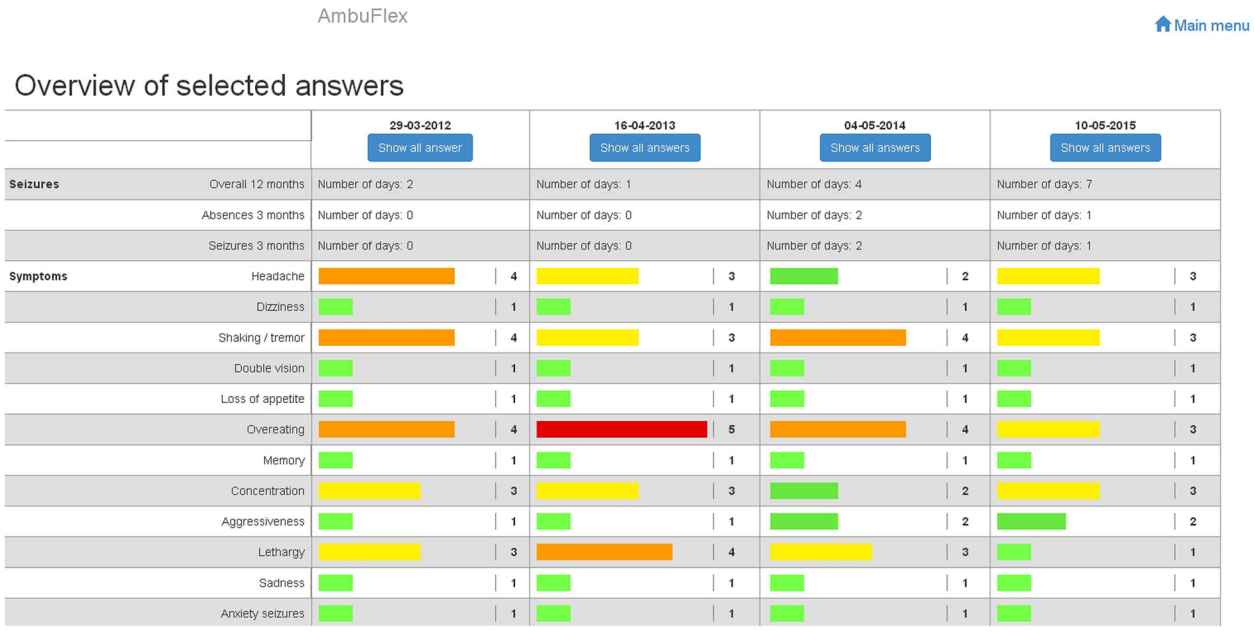


Fig. 3 PRO response overview customised to outpatients with epilepsy. A red or orange bar indicates a self reported problem, a yellow bar some problem, and green bar indicates no problems. Note: Labels were translated from Danish

c) Info: This element includes information about the open access approach including detailed information about the purpose and how to use the website. In addition, there is information about the questionnaire and why it is important to gather information about the included aspects, e.g., seizures, alcohol, pregnancy, sexuality, etc. The information provided is compiled by clinicians from the epilepsy clinics at Aarhus University Hospital and is based on disease-specific guidelines and information from the Danish Epilepsy Association.

d) Contact: Patients are asked to contact the epilepsy clinic by telephone in the event of a pressing need of attention. This element contains contact information (telephone number, email and mail addresses) to the epilepsy clinic. The emergency service is always open for those in acute need for help, for example, if the patient gets a seizure.

Randomisation

Pre-randomisation designs prevent change in behaviour in the control group because of disappointment about the allocation [19]. Eligible standard telePRO participants will be pre-randomised to standard telePRO follow-up (no change) or open access telePRO follow-up. Control as well as intervention participants receive the baseline questionnaire together with the fixed PRO questionnaire. The clinicians respond to the fixed PROs as usual and will not have access to the baseline questionnaire. Control participants will continue with fixed interval questionnaires and no change will be undertaken. Intervention participants will receive detailed information about the open access approach two weeks after a clinician's response to the fixed PRO questionnaire. The study coordinator will forward written information to the included intervention participants. Individuals who not agree to participate will continue with standard telePRO follow-up. Due to the nature of the intervention neither patients nor clinicians can be blinded to allocation. The randomisation is performed with an algorithm developed as part of the WestChronic software [12]. The allocation ratio open access/standard is 0.55/0.45. This ratio was selected to account for an expected number of patients in the open access arm who do not wish to participate.

Study timeline

Inclusion and randomisation with baseline assessments will take place from January 2016. Follow-up assessment will take place 18 months after randomisation. Baseline and follow-up assessments are shown in Table 1. Figure 4 presents the inclusion of patients and the stages in the study.

Outcomes

The effects of patient-initiated follow-up (open access telePRO) will be evaluated with regard to three different aspects: resource utilisation, quality of care, and the patient perspective. Resource utilisation will constitute the primary outcome, measured by number of contacts. Quality of care and the patient perspective constitute the secondary outcomes. Quality of care includes pivotal clinical quality measures (mortality, seizure, and treatment side effects) as well as more general patient-oriented quality measures (well-being and general health). The patient perspective includes measures related to self-management, such as health literacy, self-efficacy, and patient activation. Measures of confidence, safety, and satisfaction will be used to describe patient experiences. The patient perspective is primarily

Table 1 Primary and secondary outcomes, data sources, and timeline for measurements

Outcomes	Data sources	Measurement/month
Resource utilisation		
1. Number of contacts	The Hospital Business Intelligence Register, Central Denmark Region	0–18
Quality of care		
2. Well-being	WHO-Five Well-being Index (WHO-5)	0, 18
3. General health	Item from The Short Form Health Survey (SF-36)	0, 18
4. Mortality	The Hospital Business Intelligence Register, Central Denmark Region	0–18
5. Number of seizures	Item from the epilepsy questionnaire, Central Denmark Region	0, 18
6. Treatment side effects	Item from the epilepsy questionnaire, Central Denmark Region	0, 18
Patient perspective [a]		
7. Health literacy	The Health Literacy Questionnaire (HLQ) sub scale 4, 6 and 9	0, 18
8. Self-efficacy	General Self-Efficacy scale (GSE)	0, 18
9. Patient activation	Items from Patient Activation Measure (PAM)	0, 18
10. Confidence, safety, and satisfaction	Items from a PREM questionnaire, Danish Cancer Society	0, 18

[a] The patient perspective is primarily explored by qualitative methods in a complementary PhD study

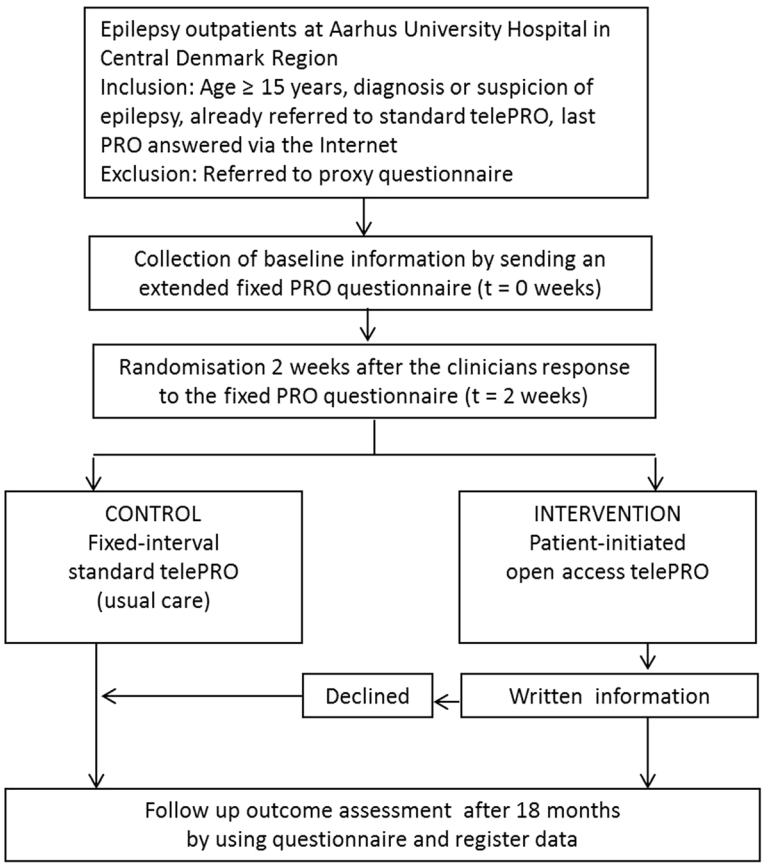


Fig. 4 Flowchart following patients from inclusion to final data collection

explored by qualitative methods in a complementary PhD study. An overview of primary and secondary outcomes, data sources, and measurement timeline is shown in Table 1.

Primary outcome

Resource utilisation

Number of contacts includes all contacts with the outpatient clinic in the study follow-up period, including face-to-face consultations with a physician, face-to-face consultations with a nurse, and telephone consultations. In addition, other health care contacts will be gathered, e.g., epilepsy-related emergency room visits and hospitalisations as well as hospitalisation related to comorbidity. Data will be gathered from the Hospital Business Intelligence Register in Central Denmark Region.

Secondary outcomes

Quality of care

Patients' well-being will be measured by using the Danish version of WHO-Five Well-being Index (WHO-5). WHO-5 was developed by the World Health Organisation for the assessment of well-being among patients with diabetes [20]. WHO-5 consists of five positively worded items reflecting present mental well-being within the previous two weeks. Items are rated on a 6-point scale ranging from 5 "all of the time" to 0 "at no time". The instrument has demonstrated sufficient psychometric properties in a wide range of chronic conditions [20, 21]. Patients' general health will be measured by using one item from the Danish version of The Short Form Health Survey (SF-36); *"In general, would you say your health is: excellent, very good, good, fair, or poor"* [22, 23]. The validity and reliability of this item are well documented [24]. Data on mortality will be gathered from the Hospital Business Intelligence Register in Central Denmark Region. Finally, number of seizures and treatment side effects will be collected from ad hoc items in the epilepsy questionnaire used at epilepsy clinics in Central Denmark Region. The validity and reliability of these items have not yet been documented.

Patient perspective

Health literacy will be measured by using the Danish version of Health Literacy Questionnaire (HLQ) [25, 26]. HLQ was developed to measure a wide range of health literacy needs of people in the community. The HLQ includes nine conceptually subscales with a total of 44

items containing five scales with agree/disagree response options and four scales with difficulties in perform tasks response options. The HLQ has well-documented psychometric properties [26]. In this study, the HLQ subscales 4, 6, and 9 will be used; *4. Social support for health, 6. Ability to actively engage with healthcare providers, 9. Understand health information well enough to know what to do.* Self-efficacy will be measured by using the Danish version of General Self-Efficacy Scale (GSE) [27, 28]. GSE was designed to assess optimistic self-belief to cope with difficult demands in life [27, 28]. GSE includes ten items with a response range from 1 "not at all true" to 4 "exactly true". The GSE scale has been used in a range of research projects in different countries and populations, where it typically yielded sufficient psychometric properties [29]. Patient activation will be measured by two ad hoc items developed with inspiration from the Danish version of the Patient Activation Measure (PAM) [30]. Confidence, safety, and satisfaction will be measured by using ad hoc items developed with inspiration from a Danish PREM (patient-reported experience measure) questionnaire from the Danish Cancer Society.

In addition, the patient perspective will be explored in a complementary qualitative PhD study. The primary aim of this study is to explore the mechanisms of actions related to standard telePRO and open access telePRO. Interpretive description (ID) will be used as the research approach [31]. Patients' experiences with telePRO will be explored in individual interviews and participant observations in outpatient clinics. The target group for participation is patients with epilepsy, referred to standard telePRO or open access telePRO follow-up in the three neurological departments in Central Denmark Region.

Other measurements

Demographic information such as sex, age, education, marital status, and duration of epilepsy diagnosis will be obtained from baseline questionnaires.

Sample size

Statistical power was estimated for the primary outcome number of contacts. Based on literature review [32] the number of consultations (n) and standard deviation (SD) was; $n = 4.64$, (SD = 2.38) in conventional follow-up and $n = 4.12$, (SD = 3.41) in open access follow-up. We expect at least a difference of one contact between the groups. Given a statistical power of 90%, p-value 0.05, and allocation ratio 0.8, we will need a sample size of 172 patients in the standard telePRO group and 214 patients in the open access telePRO group. To account for attrition and loss to follow-up, we will recruit a total of approximately 500 participants. For qualitative data, a purposeful sample of at least five participants from each group will be interviewed.

Analyses

Descriptive statistics will be used to describe differences in the baseline characteristics of participating patients in the two arms of the trial. Statistical analysis will be intention to treat, whereby all randomised participants will be included in the analysis according to their randomised allocation. The primary outcome, total *number of contacts* in the two arms, will be analysed using a sample t-test. If the distribution of data is skewed, we will use medians and nonparametric tests. For secondary outcomes, a chi-square test or logistic regression will be used for dichotomous outcome data and sample t-test or multiple linear regression analysis will be used for continuous outcome data. Non-parametric tests will be used if continuous data are not normally distributed. Demographics covariates (sex, age, education, marital status, and epilepsy diagnosis duration) will be included in the per- protocol analysis.

For qualitative data, ID will be employed as the overriding research approach. ID is an inductive research strategy in which constant comparative method with concurrent data collection and analysis is utilised to gain a deeper insight and understanding of human experiences within their natural context. The result is a comprehensive interpretation, potentially a *model of explanation* of the phenomenon under study, which can provide clinical practice with a research-based choice of action [31]. ID is considered appropriate in the present study because the approach is suited for exploration of specific clinical issues, in this case how patients with epilepsy experience standard and open access telePRO follow-up.

Ethics

The risks to participants are considered to be minimal as all eligible participants are referred to standard telePRO follow-up by clinicians at the epilepsy clinic. As a "safety net" to ensure that no patients are lost in the open access arm, the patients have to answer the epilepsy questionnaire before twice the fixed interval has elapsed. If lack of response the patient is reallocated into standard telePRO with a red status and a clinician will contact the patient. Furthermore, all patients are informed to call the clinic in pressing need of attention.

The Danish Data Protection Agency has accepted the study. In addition, the Danish research ethics committee in Central Denmark Region was contacted and has stated that approval from the committee is not necessary for this present study. Therefore, written informed consent was not obtained from the participants. Prior to study participation patients in the intervention group receive written information about the study. Study participation is entirely voluntary and participants are informed they can withdraw from the study at any time without affecting future care. In the qualitative complementary PhD study,

the participants gave written informed consent prior to enrolment, and the study was approved by the Danish Data Protection Agency.

Data security

All data activities in the study are documented and stored in the WestChronic web-system [12]. The system is physically located in Central Denmark Regions Server Park behind the firewall and Threat Management Gateway. Regular backup is performed weekly. All data transactions fulfill conditions established by the Danish Data Protection Agency.

Discussion

During the last decade, the use of PRO in clinical practice has become increasingly common, and to our knowledge, AmbuFlex is the first generic PRO system that uses PRO as the basis for outpatient follow-up [13]. The focus of this trial will be to evaluate the effect of a patient-initiated open access telePRO intervention compared to standard telePRO with respect to resource utilisation, quality of care, and the patient perspective. Ideally, we would have preferred to compare the two arms (standard and open access) with conventional follow-up with pre-booked outpatient visits to the clinic. However, this was not possible because the epilepsy clinics in Central Denmark Region have used standard telePRO follow-up since 2012. Thus, we will compare two rather similar outpatient follow-up activities, which will probably result in only small differences in effect between the groups. Evaluation of the effect must be done using reliable, valid, and clinically meaningful measures. This study includes outcome measures based on recommendations from clinical experts, researchers, and the literature [33].

Loss to follow-up is one of the main concerns in randomised controlled studies [34]. Loss to follow-up in this study is related to the open access group of patients since study participation in the open access arm is entirely voluntary, and participants can choose to continue with standard telePRO follow-up. Loss of statistical efficiency can be overcome by increased the number of participants in the study [19]. We have taken this into consideration and will include 10% more patients in the open access telePRO arm. In addition, we will recruit a larger number of participants than the minimum sample size calculation indicated.

Only web-responders will be included in the open access arm, and the results may therefore be generalizable only to this subgroup of epilepsy patients. These patients may differ with respect to education, age, and use of new technologies compared to the entire group of epilepsy patients. In another study in progress, the aim is to examine determinants for referral to telePRO follow-up. Data from this study can be used to compare the study population with the entire group of patients with epilepsy.

Another potential challenge may be how individuals in the intervention group will use the "My Epilepsy" website. Some patients are better able to decide themselves when they need to contact the clinic, while others are more reserved and afraid to be a nuisance. Several patients that have used standard telePRO have pointed out the benefit of getting a fixed questionnaire once a year. They do not believe they would remember to answer if they had to do it on their own. This could signify that even though they may not feel the need for a clinical appointment, but do feel a form of security in answering the fixed interval questionnaire. This will be taken into consideration in the study, since all patients in the intervention group will receive a questionnaire if they do not respond within two times the referred interval, for example, within 24 months if they are assigned a 12-month questionnaire interval. Another concern could be that patients in the open access group could choose to make a call instead of answering the questionnaire when they need to get in contact with the clinic. If they behave in this way, the benefit of using PRO in clinical practice will be reduced.

Standard telePRO has been well integrated into clinical practice in three epilepsy clinics in Central Denmark Region since 2012. A new patient-initiated approach has been developed that may result in potential benefits in terms of the patient perspective and resource utilisation. The potential benefits as well as possible drawbacks need to be evaluated. We have decided to combine qualitative and quantitative research methods in two parallel PhD studies. The intention of the complementary qualitative PhD study is to further explain the findings from the randomised study by providing a description of the various ways in which telePRO is manifested and an interpretation of the underlying mechanisms of action. The two studies will complement each other and contribute with important research-based knowledge to guide future telePRO interventions in relation to effect on resource utilisation, quality of care, and the patient perspective.

Trial status

On going.

Additional file

Additional file 1: SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*. (DOC 122 kb)

Abbreviations

GSE: General Self-Efficacy scale; HLQ: Health literacy questionnaire; ID: Interpretive description; PAM: Patient activation measure; PREM: Patient-reported experience measure; PRO: Patient-reported outcome; SD: Standard deviation; SF-36: Short form health survey; WHO-5: WHO-Five well-being index

Acknowledgements
This study is funded by Aarhus University, the Health Research Fund of Central Denmark Region, and the Danish foundation TrygFonden.

Funding
Aarhus University, the Health Research Fund of Central Denmark Region, and the Danish foundation TrygFonden have funded this PhD study. The funding sources had no role in design of this study and will not have any role during its execution, analysis, interpretation of the data, or decision to submit results.

Availability of data and materials
Not applicable.

Authors' contributions
NHH and LMVS conceived the study in collaboration with PS, AdT, and KL. LMVS, KHP, CTM, and NHH participated in the development of the open access intervention. KHP and NHH developed the open access website interface and the randomisation algorithm as part of the WestChronic software. LMVS participated in recruitment of participants, data collection and registration. LMVS drafted the manuscript. All authors have read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
The Danish research ethics committee in Central Denmark Region was contacted. According to Danish law, the committee has stated that approval from the committee is not necessary for this present study. Therefore, written informed consent was not obtained from the participants.

Data monitoring and dissemination policy
A data monitoring committee was not needed due to minimal risk in the intervention group. The results of the study will only be published in peer reviewed journals.

Protocol version
Issue date 6 Jan 2017, version number: 01.

Author details
[1]AmbuFlex, Regional Hospital West Jutland, Herning, Denmark. [2]The Research Programme in Patient involvement, Aarhus University Hospital, Aarhus, Denmark. [3]Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark. [4]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark. [5]Department of Neurology, Aarhus University Hospital, Aarhus, Denmark. [6]Department of Public Health, Aarhus University, Aarhus, Denmark. [7]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark.

Received: 15 November 2016 Accepted: 13 January 2017
Published online: 26 January 2017

References

1. Indenrigs- og Sundhedsministeriet. Analyse af aktiviteten i sygehusvæsenet. 2010. http://www.sum.dk/~/media/Filer%20-%20Publikationer_i_pdf/2010/Analyse_aktiviteten_sygehusv/Analyse_sygehusv.ashx. Accessed 6 June 2016.
2. Whear R, Abdul-Rahman AK, Boddy K, Thompson-Coon J, Perry M, Stein K. The clinical effectiveness of patient initiated clinics for patients with chronic or recurrent conditions managed in secondary care: a systematic review. PLoS One. 2013;8(10):e74774.
3. US Department of Health and Human Services Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. 2009. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Accessed 6 Jun 2016.
4. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J Eval Clin Pract. 2006; 12(5):559–68.
5. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. BMC Health Serv Res. 2013;13: 211–6963. 13-211.
6. Johansen MA, Berntsen GK, Schuster T, Henriksen E, Horsch A. Electronic symptom reporting between patient and provider for improved health care service quality: a systematic review of randomized controlled trials. part 2: methodological quality and effects. J Med Internet Res. 2012;14(5):e126.
7. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, Revicki DA, Symonds T, Parada A, Alonso J. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res. 2008;17(2):179–93.
8. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. CA Cancer J Clin. 2012;62(5):336–47.
9. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. BMJ Qual Saf. 2014;23(6):508–18.
10. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? Qual Life Res. 2009;18(1):115–23.
11. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns. 2002;48(2):177–87.
12. Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjaer E, Schougaard LM. Use of Patient-Reported Outcome (PRO) measures at group and patient levels: experiences from the generic integrated PRO system, WestChronic Interact. J Med Res. 2014;3(1):e5.
13. Schougaard LM, Larsen LP, Jessen A, Sidenius P, Dorflinger L, de Thurah A, Hjollund NH: AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Qual Life Res. 2016; 25(3):525–34.
14. Danish Agency for Digitalisation and PA Consulting Group. Analyse af Patient Rapporterede Oplysninger (PRO). 2015. http://www.digst.dk/Digital-velfaerd/Strategi-for-digital-velfaerd/Rapporter-og-analyser/Sundhed/Analyse-af-patientrapporterede-oplysninger. Accessed 6 Jun 2016.
15. Center for Telemedicin, Region Midtjylland. AmbuFlex – Behovsstyrede ambulante patientforløb med spørgeskemabaseret klinisk selvmonitorering. Bilag: Patientperspektiv – interview med 13 epilepsipatienter. 2013 (In Danish).
16. Taneja A, Su'a B, Hill AG. Efficacy of patient-initiated follow-up clinics in secondary care: a systematic review. Intern Med J. 2014;44(12a):1156–60.
17. Whear R, Abdul-Rahman AK, Thompson-Coon J, Boddy K, Perry MG, Stein K. Patient initiated clinics for patients with chronic or recurrent conditions managed in secondary care: a systematic review of patient reported outcomes and patient and clinician satisfaction. BMC Health Serv Res. 2013; 13:501. 6963-13-501.
18. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleza-Jeric K, Hrobjartsson A, Mann H, Dickersin K, Berlin JA, Dore CJ, Parulekar WR, Summerskill WS, Groves T, Schulz KF, Sox HC, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.
19. Zelen M. A new design for randomized clinical trials. N Engl J Med. 1979; 300(22):1242–5.
20. Bech P, Gudex C, Johansen KS. The WHO (Ten) well-being index: validation in diabetes. Psychother Psychosom. 1996;65(4):183–90.
21. Topp CW, Ostergaard SD, Sondergaard S, Bech P. The WHO-5 well-being index: a systematic review of the literature. Psychother Psychosom. 2015; 84(3):167–76.
22. Bjorner JB, Damsgaard MT, Watt T, Groenvold M. Tests of data quality, scaling assumptions, and reliability of the Danish SF-36. J Clin Epidemiol. 1998;51(11):1001–11.
23. Bjorner JB, Thunedborg K, Kristensen TS, Modvig J, Bech P. The Danish SF-36 Health Survey: translation and preliminary validity studies. J Clin Epidemiol. 1998;51(11):991–9.
24. Gandek B, Ware Jr JE, Aaronson NK, Alonso J, Apolone G, Bjorner J, Brazier J, Bullinger M, Fukuhara S, Kaasa S, Leplege A, Sullivan M. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol. 1998;51(11):1149–58.

25. Maindal HT, Kayser L, Norgaard O, Bo A, Elsworth GR, Osborne RH. Cultural adaptation and validation of the Health Literacy Questionnaire (HLQ): robust nine-dimension Danish language confirmatory factor model. Springerplus. 2016;5(1):1232. 016-2887-9 . eCollection 2016.
26. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health. 2013;13:658. 2458-13-658.
27. Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol. 2005;139(5):439–57.
28. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in health psychology: a user's portfolio. causal and control beliefs. Windsor: NFER-NELSON; 1995. p. 35–7.
29. Scholz U, Gutiérrez-Doña B, Sud S, Schwarzer R. Is general self-efficacy a universal construct? Psychometric findings from 25 countries. Eur J Psychol Assess. 2002;18(3):242–51.
30. Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form patient activation measure (PAM13) in a Danish version. BMC Public Health. 2009;9:209. 2458-9-209.
31. Thorne S. Interpretive Description. Walnut Creek: Left Coast Press, Inc; 2008.
32. Williams JG, Cheung WY, Russell IT, Cohen DR, Longo M, Lervy B. Open access follow up for inflammatory bowel disease: pragmatic randomised trial and cost effectiveness study. BMJ. 2000;320(7234):544–8.
33. Fountain NB, Van Ness PC, Swain-Eng R, Tonn S, Bever Jr CT, American Academy of Neurology Epilepsy Measure Development Panel and the American Medical Association-Convened Physician Consortium for Performance Improvement Independent Measure Development Process. Quality improvement in neurology: AAN epilepsy quality measures: report of the quality measurement and reporting subcommittee of the American academy of neurology. Neurology. 2011;76(1):94–9.
34. Sorensen HT, Lash TL, Rothman KJ. Beyond randomized controlled trials: a critical comparison of trials with nonrandomized studies. Hepatology. 2006; 44(5):1075–82.

SPIRIT

STANDARD PROTOCOL ITEMS: RECOMMENDATIONS FOR INTERVENTIONAL TRIALS

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	_____1_____
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	_____1_____
	2b	All items from the World Health Organization Trial Registration Data Set	_____1_____
Protocol version	3	Date and version identifier	_____10_____
Funding	4	Sources and types of financial, material, and other support	_____10__
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	_____1+10_____
	5b	Name and contact information for the trial sponsor	_____10_____
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	_____10_____
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	_____

Introduction

Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	___2-3_____
	6b	Explanation for choice of comparators	___2-3_____
Objectives	7	Specific objectives or hypotheses	____3_____
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	____3_____

Methods: Participants, interventions, and outcomes

Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	____3_____
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	____3_____
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	___3-6_____
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	_____6_____
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	_____6_____
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	_____
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	__6-8 + Table 1
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	__6 + Figure 4_

Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	___8_____
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	___8_____

Methods: Assignment of interventions (for controlled trials)

Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	____6_____
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	____6_____
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	____6 + 10____
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	____6_____
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	_____

Methods: Data collection, management, and analysis

Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	_7-8 + Table 1
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	__6 + 9_____

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	9
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	8
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	8
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	10
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	8
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	10
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	_____
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	_____
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	___9_____
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	___10_____
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	___10_____
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	_____
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	_____10_____
	31b	Authorship eligibility guidelines and any intended use of professional writers	_____
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	_____
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	__10_____
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	_____

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

PAPER V

Including online supplemental materials:

Additional file 1: Table 1. Information for Reporting Randomized Controlled Trials With Patient reported Outcomes

Additional file 2: Note: The epilepsy questionnaire can be found in the Appendices section (Appendix II)

Additional file 3: Supplemental analyses (stratified analyses and sensitivity analyses) and the randomization computer code.

Journal of Patient-Reported Outcomes

RESEARCH Open Access

Patient-initiated versus fixed-interval patient-reported outcome-based follow-up in outpatients with epilepsy: a pragmatic randomized controlled trial

Liv Marit Valen Schougaard[1*], Caroline Trillingsgaard Mejdahl[1], Jakob Christensen[2,3], Kirsten Lomborg[3], Helle Terkildsen Maindal[4], Annette de Thurah[3,5] and Niels Henrik Hjollund[1,6]

Abstract

Background: The use of patient-reported outcome (PRO) could potentially contribute to the reorganization of the health care system. AmbuFlex is a PRO system used in remote patient monitoring, in which questionnaires are sent to patients at fixed intervals. The PRO data are used by clinicians to decide whether patients need clinical attention. Better self-management and cost-saving follow-up activities may be achieved by letting patients initiate need of contact. We evaluated the effects of patient-initiated PRO-based outpatient follow-up on health care resource utilization, quality of care, and the patient perspective.

Methods: We conducted a parallel two-arm pragmatic randomized controlled trial at the Department of Neurology, Aarhus University Hospital, Denmark. Outpatients with epilepsy (≥ 15 years old), attending fixed-interval PRO-based follow-up with web-based questionnaires, were randomly assigned in a ratio of 0.55:0.45 to either 1) patient-initiated PRO-based follow-up (open access telePRO) or 2) fixed-interval PRO-based follow-up (standard telePRO). The primary outcome was the number of outpatient hospital contacts related to epilepsy retrieved from a regional registry. Hospitals admissions and emergency room visits were also assessed. Secondary self-reported outcomes including general health, well-being, health literacy, self-efficacy, number of seizures, side effects, confidence, safety, and satisfaction were retrieved from questionnaires. Data were analyzed by the intention-to-treat and per-protocol approaches.

Results: Between January 2016 and July 2016, 593 patients were randomized to either open access telePRO ($n = 346$) or standard telePRO ($n = 247$). At 18 months, no statistically significant differences were found between the arms regarding number of telephone consultations or outpatient visits. Patients in the open access arm had a slightly lower, statistically significant number of emergency room visits than patients in the standard arm. Self-reported mental well-being in the open access arm was slightly, statistically significantly lower than in the standard arm. Other secondary outcomes did not differ statistically significantly between arms.

Conclusion: This study did not find, as hypothesized, less use of health care resources or improved patient self-management or satisfaction in the patient-initiated PRO-based initiative compared to fixed-interval PRO-based follow-up. Patient-initiated PRO-based follow-up may be used as an alternative to fixed-interval PRO-based follow-up in patients who prefer this approach, but there is insufficient evidence for recommending a system-wide shift to patient-initiated PRO-based follow-up.

Trial registration: Registered 4 February 2016 with ClinicalTrials.gov: NCT02673580.

Keywords: Patient reported outcome measures, Randomized controlled trial, Ambulatory care, Outpatient clinics, hospital, Epilepsy

* Correspondence: livschou@rm.dk
[1]AmbuFlex/WestChronic, Occupational Medicine, University Research Clinic, Aarhus University, Gl. Landevej 61, DK-7400 Herning, Denmark
Full list of author information is available at the end of the article

Introduction

Health care systems are experiencing an increased volume of patients with chronic conditions concurrent with increased focus on patient involvement and patient self-management [1, 2]. The use of patient-reported outcome (PRO) measures in clinical practice could potentially contribute to reorganization of the health care system and support patient involvement. PRO is a measurement directly reported by the patients based on their own perceived symptoms and health status [3]. Santana et al. describe a theoretical framework outlining the potential effects of using PRO measures in the care of chronically ill patients [2]. According to this framework, letting patients contribute with self-reported information about the impact of their disease and its treatment can contribute to better communication, engagement, self-management, and patient outcomes [2]. A number of reviews including randomized controlled studies have found that the use of PRO measures in a clinical setting improved patient-clinician communication, patient satisfaction, and detection of patients' functional and mental health problems [4–6]. Furthermore, PRO has the potential to inform clinical decision-making and support self-management [7–9]. Findings related to clinical patient outcomes are less consistent [4, 6].

In Denmark, a generic configurable PRO solution, AmbuFlex, uses PRO measures as the very basis for outpatient follow-up in several chronic and malignant diseases [8, 10] including follow-up for outpatients with epilepsy. Epilepsy is a chronic condition characterized by recurrent seizures affecting functional, mental, and social aspect of life [11, 12]. Studies have reported that persons with seizures have increased risk of mood disorders, reduced quality of life, and significantly more social stigma than persons with no seizures [12, 13]. These findings support the need for differential and individualized follow-up in the care for patients with epilepsy. Several PRO measures have been developed for use in patients with epilepsy at the aggregated level [14]; however, the evidence regarding use of PRO measures on the individual level is weak [15, 16]. In 2012, an epilepsy version of AmbuFlex was developed in Central Denmark Region, here called standard telePRO [8]. In standard telePRO, the patients receive fixed-interval questionnaires at home instead of having pre-scheduled appointments at the outpatient clinic. Based on an automated algorithm, the patients' PRO measures are used to decide the need for clinical attention, potentially leading to fewer visits and thus less treatment burden for well-treated patients. If the patient needs attention, the PRO measures are used to support patient-clinician communication.

Standard telePRO may not be an adequate solution if the patients have a variable need of clinical attention. A more patient-centered solution based on patient preferences to decide the timing of a clinical contact may be even more beneficial to enhance patient involvement and management of own care. Several reviews of randomized controlled trials have investigated the effect of patient-initiated interventions in which patients have direct access to the outpatient clinic if needed [17–19]. In studies of patients with rheumatoid arthritis, inflammatory bowel disease, and breast cancer, no differences were seen regarding clinical- or patient-reported health outcomes between patient-initiated intervention groups and clinician-initiated control groups. Furthermore, some studies found higher patient satisfaction and lower use of health care resources in the patient-initiated model [17–19]. We have not been able to find any studies that use a patient-initiated model in patients with epilepsy and no studies that use PRO as the main access point to flag the need for clinical attention.

In this study we evaluated the effects of patient-initiated outpatient follow-up in patients with epilepsy. The specific aims were to compare utilization of health care resources, quality of care, and the patient perspective in two outpatient follow-up activities: patient-initiated PRO-based follow-up (open access telePRO) versus fixed-interval PRO-based follow-up (standard telePRO).

We hypothesized that the number of contacts would be lower, quality of care at least as good, patient self-management better, and patient evaluation of health service improved among patients in the open access telePRO arm compared with those in the standard telePRO arm.

Methods

Study design

This study was a parallel two-arm pragmatic randomized controlled trial in which the participants were allocated to either patient-initiated PRO-based follow-up (open access telePRO) or fixed-interval PRO-based follow-up (standard telePRO). The study was carried out among epilepsy outpatients at the Department of Neurology, Aarhus University Hospital in Denmark. Standard telePRO has been used at the department since 2012. In January 2016, approximately 2500 epilepsy outpatients were attending standard telePRO follow-up. The study followed the Consolidated Standard of Reporting Trial (CONSORT) guideline for reporting parallel group randomized trials [20] and the CONSORT PRO extension [21] (Additional file 1). A study protocol has been published [22].

Participants and settings

Participants were included between January 2016 and July 2016. From January 2016, all patients in standard telePRO follow-up at the Department of Neurology, Aarhus University Hospital, received a baseline research questionnaire combined with the fixed-interval epilepsy questionnaire from the outpatient clinic. Patients could choose to respond via a paper or web version of the questionnaires.

Clinicians assessed the fixed-interval epilepsy questionnaire according to their normal routine, but were blinded to the research questionnaire. Approximately 14 days after the patients filled in the research questionnaire, eligible participants were randomized to either open access telePRO or standard telePRO. Patients were eligible if they were ≥ 15 years old, had an epilepsy diagnosis or suspicion of epilepsy, were attending standard telePRO follow-up, and had filled in the last questionnaire via the Internet. Patients were excluded if they were paper respondents or if they had stop attending standard telePRO follow-up before randomization.

The study coordinator enrolled participants approximately once a week during the inclusion period. After randomization, participants in the open access arm received detailed written information about the intervention via surface mail sent by the study coordinator. Participants were requested to contact the study coordinator if they did not want to participate in the open access telePRO intervention and preferred to continue with standard care (standard telePRO). Standard arm participants continued standard telePRO and there was no change in the follow-up. Blinding of the randomization allocation was not possible for either participants or clinicians. Follow-up assessments were conducted approximately 18 months after randomization [22]. The rationale for 18 months was based on the fact that more than in 80% of patients attending standard telePRO follow-up, questionnaires were sent at fixed 12-months intervals. This means that patients in standard telePRO follow-up may not have had contact with the outpatient clinic before 12 months had passed; thus, a follow-up period longer than 12 months was required.

Pre-randomization

According to the inclusion criteria, patients were pre-randomized [23] in a ratio of 0.55:0.45 to either open access telePRO or standard telePRO. In a pre-randomization design, patients in the intervention arm are informed about the allocation following randomization, and disappointment about the allocation in the control arm can be prevented [23]. The skewed randomization allocation was applied because of an expected higher number of drop outs in the open access arm compared to the standard arm [22]. A higher dropout rate in the open access arm was expected, since participation was voluntary and the participants could at any time during the study decide to continue standard telePRO, if, for example, they did not want to initiate contact to the clinic by themselves, but rather receive questionnaires at fixed intervals. To account for this, we decided to randomize 10% more patients to the open access arm, as this would enhance the statistical power of the per-protocol analysis [24]. We used simple randomization due to an expected large study population and did not block randomization or other procedures to help achieve balance in the number or characteristics of the participants in the two arms. Computer-generated randomization was used. The computer code was developed and integrated into the WestChronic/ AmbuFlex system (Additional file 3, page 11) [10].

Interventions

Standard arm – standard telePRO (usual care)

In standard telePRO, patients filled in fixed-interval disease-specific questionnaires every 3, 6, or 12 months, which were used as a partly automatic tool to support the decision regarding whether the patient needed clinical attention at the present time [8]. In the questionnaire, all patients could request a telephone consultation or an appointment in the outpatient clinic, regardless of their response to the other questions in the questionnaire. The questionnaire development is described elsewhere [25], and the questionnaire can be found in the Additional file 2.

The patient's response to the questionnaires was given a green, yellow, or red color by using a pre-defined automated algorithm [8, 25]. Green indicated no need of clinical attention, red indicated need of attention, whereas yellow indicated that the patient might need attention. Green responses were handled automatically by the server software, and a new questionnaire was automatically scheduled to be sent to the patient at the pre-defined fixed interval, for example, after 12 months. All yellow and red responses were shown on an alert list, available to the clinicians, who accessed the list daily. A red response indicated need of clinical attention, and the clinician contacted the patient as quickly as possible. Patients were either contacted by telephone or they received a face-to-face appointment. For yellow responses, patients were only contacted if the clinicians judged that it was necessary. The patient's questionnaire response was graphically presented to the clinicians, who accessed all the yellow and red responses through the Electronic Health Record system together with other relevant data from the record (laboratory tests, medication, etc.) [8, 22].

Intervention arm – open access telePRO

For patients randomized to open access telePRO, patient contact with the outpatient clinic was based on the patient's preferences. Patients were asked to contact the outpatient clinic by themselves when they felt it necessary. Thus, at any time during the follow-up period, these patients could indicate a need for contact with the outpatient clinic by filling in the disease-specific questionnaire (Additional file 2). For this purpose, an open access website 'My Epilepsy' was developed. The website contains four core elements to allow patients to: 1) answer a questionnaire when they needed to get in contact with the clinic, 2) view their previously questionnaire responses, 3) view information about the questionnaire,

and 4) view contact information (e.g. telephone number) to the outpatient clinic [22]. Full detail of the development and features of this website are available elsewhere [22]. Patients had access to the open access website via a secure login to the Danish ehealth Portal "Sundhed.dk". In addition, the patients could also phone the outpatient clinic if needed. All questionnaire responses in the open access arm turned red (definite need of attention) on the alert list to the clinicians, since these patients were instructed to only fill in the questionnaire if they needed to talk to a clinician. The clinician checked the alert list daily and assessed the red open access responses as quickly as possible in the same web-system as in standard telePRO [8, 22]. The patients were contacted by telephone, and a face-to-face appointment was scheduled if necessary. If the patient did not fill in a questionnaire to the outpatient clinic within a priori defined time-span, the web system automatically sent a reminder to the patients with instructions to fill in the questionnaire. For example, a reminder was sent after 12 months if the patient prior to randomization was originally referred to a 6-month fixed questionnaire interval in standard telePRO. The clinicians also received information on the alert list about patients who did not respond to these reminders, and they were subsequently contacted by a clinician.

Outcomes

Primary outcome

The primary outcome was the number of outpatient hospital contacts related to epilepsy from baseline to follow-up (timeframe 18 months). The number of contacts included all outpatient telephone consultations and outpatient visits (face-to-face consultations) with a nurse or a physician. Data regarding hospital admissions and emergency room visits were also assessed. The number of telephone consultations, outpatient visits, hospital admissions, and emergency room visits during the 18-month period were retrieved separately from a regional registry: the Business Intelligence Register in Central Denmark Region, which contains information about routinely collected activity measures from the Department of Neurology and Aarhus University Hospital.

Secondary outcomes

Secondary self-reported outcomes were retrieved from the baseline research questionnaire and a follow-up research questionnaire sent to patients before randomization and 18 months after randomization. Both questionnaires included information about number of seizures, side effects, well-being, general health, health literacy, self-efficacy, patient activation, confidence, safety, and satisfaction.

Clinical outcome measures The number of seizures last year and the degree of side effects were extracted from two single items in the epilepsy questionnaire (Additional file 2). Test-retest reliability of the side effects item has been reported to be substantial [25], but validity has not yet been reported. The side effects item ranges from 1 (best) to 4. Mortality was recorded at the end of the follow-up period and retrieved from the Business Intelligence Register in Central Denmark Region.

Patient-centered outcome measures Well-being was extracted from the WHO-5 Well-Being Index (WHO-5) [26, 27]. WHO-5 is a generic questionnaire, and the psychometric findings have been reported in other patient populations [27]. The WHO-5 includes five items which are used to calculate a score that ranges from 0 (worst) to 100. General health (GH) was extracted from one single item: "In general, would you say your health is: excellent, very good, good, fair, or poor" from the generic questionnaire: The Short Form Health Survey SF-36 [28, 29]. The GH item was scaled from 1 (best) to 5.

Patient self-management Health literacy was extracted from the generic Health Literacy Questionnaire (HLQ), sub-scale 4: "Social support for health", sub-scale 6: "Ability to actively engage with healthcare providers", and sub-scale 9: "Understand health information well enough to know what to do" [30, 31]. HLQ sub-scale 4 is a 4-item scale that ranges from 1 (worst) to 4, whereas HLQ sub-scales 6 and 9 are 5-item scales ranging from 1 (worst) to 5. Self-efficacy was extracted from the generic 10-item General Self-Efficacy Scale (GSE) [32, 33]. The psychometric properties of GSE have been evaluated across many countries [32]. The GSE score ranges from 10 (worst) to 40. Patient activation was extracted from two single items modified from a generic questionnaire: the Patient Activation Measure (PAM) [34]. The two PAM items: "I am confident that I can tell when I need to get outpatient care" and "I am confident I can figure out solutions when new situations or problems arise with my health condition" range from 1 (worst) to 4.

Patient health service evaluation Confidence, safety, and satisfaction were extracted from three single items, which were modified from a patient-reported experiences questionnaire developed by the Danish Cancer Society [35]. Psychometric properties have not been reported. Scores for the three items range from 1 (best) to 4.

Other measurements All Danish Citizens have a 10-digit unique personal identification number assigned to all citizens at birth [36]. It encodes gender and date of birth, and was used to calculate age and gender at baseline. Other patient characteristics were extracted from the baseline research questionnaire including cohabitation status, education, and duration of epilepsy. The

education variable was categorized into three levels: no or low (primary and lower secondary school), medium (upper secondary school and short cycle tertiary), and high (bachelor and master). Duration of epilepsy was divided into two groups with a cut-off point at 2 years duration, as this was considered an acceptable level of being experienced or not-experienced with the epilepsy diagnosis.

Process evaluation Automated computer logs were used to track and evaluate use of the open access website "My Epilepsy" in the WestChronic/AmbuFlex-system [10]. Use was defined as number of questionnaires filled in by the patients. Number of reminders mailed to patients and the number of patients who responded to these reminders were also logged into the system.

Sample size

Based on a two-sided statistical test, the study was designed to have a power of 90% (*P*-value 0.05) [22]. This was based on a study that reported that the mean number and standard deviations (SDs) of outpatient visits were 4.12 (SD = 3.41) in the open access arm and 4.64 (SD = 2.38) in the control arm [37]. We expected to detect a difference of at least one contact between the arms. This required a sample size of 386 participants. To account for attrition and loss to follow-up, the sample size was supplemented with 207 patients (132 in the open access arm and 75 in the control arm).

Statistical methods

Data were analyzed based on the intention-to-treat (ITT) approach. Between-arm differences in the number of outpatient visits, telephone consultations, hospital admissions, and emergency room visits were analyzed by simple linear regression. Because the normality distributions were skewed, 95% confidence intervals were found by using the bootstrap method with 1000 replications [38, 39]. Between-arm differences in all secondary outcomes, apart from mortality, were analyzed by multiple linear regression by calculating differences at follow-up (18 months) adjusted for the baseline value.

Between arm differences were also analyzed on a per-protocol basis. The per-protocol analysis included only participants who completed the open access intervention. Patients were defined as 'completers' if they did not decline to participate in the intervention during the study period. Between-arm differences in the number of outpatient visits, telephone consultations, hospital admissions, and emergency room visits were analyzed by multiple linear regression adjusted for gender, age, education, cohabitation status, epilepsy duration, and seizures during last year. Confidence intervals were found by using the bootstrap method with 1000 replications [38, 39]. Between-arm differences at follow-up of secondary outcomes were analyzed by multiple linear regression adjusted for the baseline value, gender, age, education, cohabitation status, epilepsy duration, and seizures during last year.

Differences in baseline data between the arms were evaluated by chi-squared test for categorical variables and the Wilcoxon Mann–Whitney test or unpaired *t*-test for continuous variables. Normally distributed baseline data were presented with means and SDs, otherwise medians and interquartile ranges (IQR) were reported additionally. In PRO measures, information about item nonresponse was presented as numbers and percentages. Estimation of sum scores followed guidelines for handling items missing for each specific score. In the HLQ scores, the mean scores of the other items were used to estimate the score. If more than two items were missing, the score was not estimated. This instruction was obtained from a user package we received after signing a license agreement. We were not able to find a standardized guideline for the other scales (GSE and WHO-5) and thus, we decided not to calculate the score if items were missing.

To explore whether the results changed in sub-groups in the population, we performed supplemental explorative ITT analyses by stratifying on age, gender, and high/low health literacy (HLQ4: "social support for health"). The median values for age (median value = 45.7 years) and the HLQ4 scale (median value = 3.4) at baseline were used to define the threshold of the high (≥ 45.7 years and ≥ 3.4 HLQ4 score) and low groups (< 45.7 years and < 3.4 HLQ4 score). Furthermore, ITT-sensitivity analyses were used to establish the impact of missing self-reported data at follow-up. Sensitivity analyses were only performed for the WHO-5 score. If the WHO-5 score was missing at follow-up, the score was imputed by using the WHO-5 score from the baseline research questionnaire. Four scenarios of the imputed follow-up values were considered: 1. the baseline value was reduced with 5 points in the open access arm and was unchanged in the standard arm, 2. the baseline value was reduced with 5 points in the standard arm and was unchanged in the open access arm, 3. the baseline value was increased with 5 points in the open access arm and was unchanged in the standard arm, and 4. the baseline value was increased with 5 point in the standard arm and was unchanged in the open access arm. Then, between-arm ITT-differences in the WHO-5 score at follow-up were analyzed by multiple linear regression adjusted for the baseline WHO-5 value. All analyses were conducted in STATA version 15 (Stata Corporation, College Station, Texas, USA).

Results

Participant flow and baseline data

A total of 593 outpatients with epilepsy were included from January 2016 to July 2016; 346 were randomized to the open access telePRO arm and 247 to the standard

telePRO arm (Fig. 1). A total of six patients died (two patients in the open access arm and four patients in the standard arm), and one patient from the open access arm moved abroad within the follow-up period of 18 months: these seven patients were not included in the analyses. With respect to secondary self-reported outcomes, 202 (58%) in the open access arm and 150 (61%) in the standard arm responded to the questionnaire at both baseline and at follow-up. During the follow-up period, 43 patients declined to participate in the open access intervention and were excluded in the per-protocol analyses. The baseline characteristics of the participants are shown in Table 1. No statistically significant baseline differences were found between the open access arm and the standard arm. Also, no statistically significant differences were found between patients who completed the open access intervention (the per-protocol arm) and the standard arm.

Primary outcomes

No statistically significant differences were found between the arms regarding mean number of telephone consultations or outpatient visits (Table 2). The mean difference in telephone consultations between the open access arm and the standard arm was − 0.32 (95% CI: − 0.68 to 0.05). Patients in the open access arm had a statistically significant, slightly lower number of emergency room visits than those in the standard arm; the mean difference was − 0.11 (95% CI: − 0.21 to − 0.01). No statistically significant difference was found in hospital admissions.

Secondary outcomes

No statistically significant differences were found between the open access arm and the standard arm regarding clinical outcome measures such as seizures during the last year, side effects (Table 3), and mortality. Patient-centered outcome measures showed a statistically significant difference of − 3.21 (95% CI: − 6.38 to − 0.05) in the WHO-5 well-being score at follow-up, giving a lower score in the open access arm than in the standard arm. General health status did not differ between the two arms. Furthermore, no statistically significant differences were found in outcome measures related to patient self-management (health literacy, self-efficacy, patient activation) and health service evaluation (confidence, safety, satisfaction).

Per-protocol and stratified analyses

Results from per-protocol analyses are shown in Tables 2 and 3. No statistically significant differences were found in either primary or secondary outcomes. Explorative stratified ITT analyses with stratification on gender and high/low health literacy did not change the results noticeably (Additional file 3, pages 1 to 9). After stratification on age in the low age group (median age below 45.7 years), the participants in the open access arm had fewer telephone consultations and emergency room visits, − 0.67

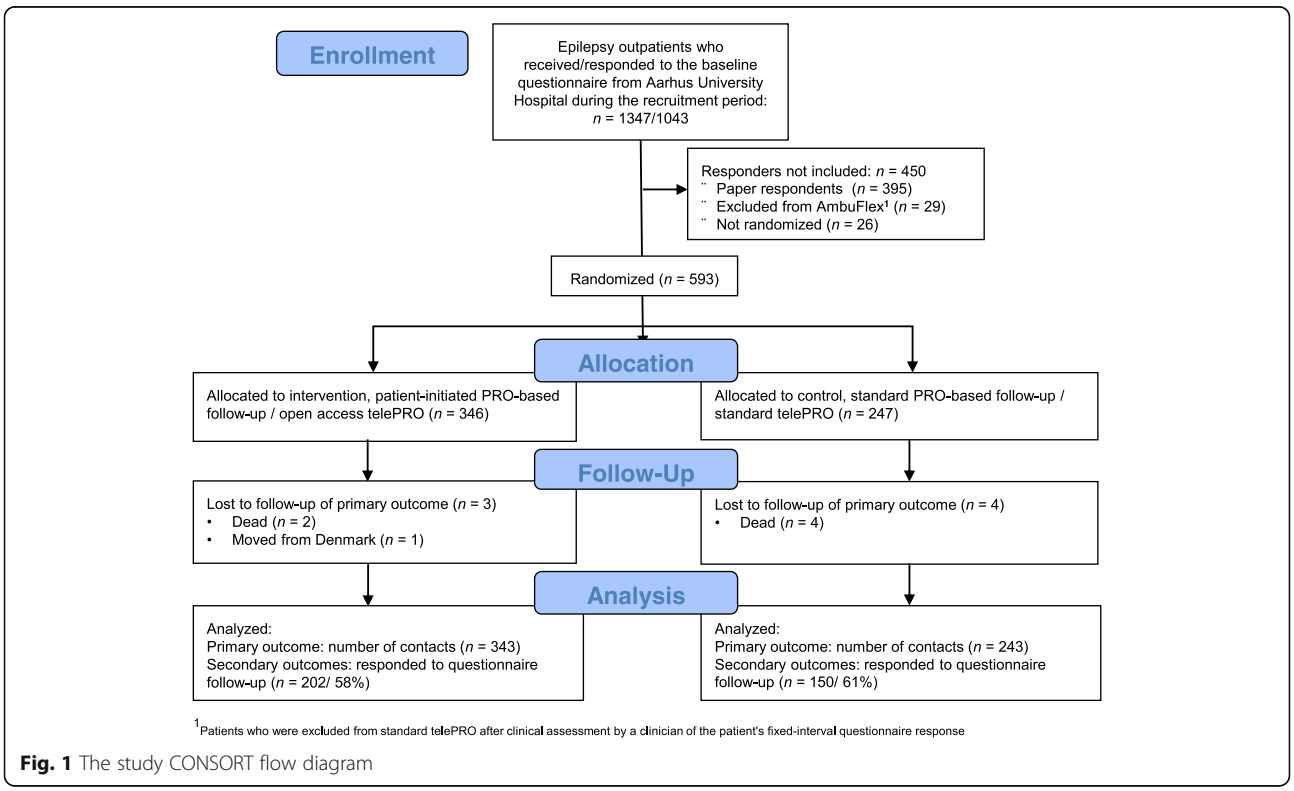


Fig. 1 The study CONSORT flow diagram

Table 1 Participant baseline characteristics, $N = 593$

Variables	Intention-to-treat population		Per-protocol population
	Intervention arm (open access telePRO): $n = 346$	Control arm (standard telePRO): $n = 247$	Patients who completed the open access intervention: $n = 300$
Mean (SD) age, years	46.3 (17.2)	47.2 (17.3)	45.8 (17.1)
Gender: male, *n* (%)	182 (53)	115 (47)	164 (55)
Cohabitation status, *n* (%)			
Living alone	76 (22)	62 (25)	59 (20)
Missing	12 (3)	3 (1)	12 (4)
Education, *n* (%)			
No or low	94 (27)	62 (25)	76 (25)
Medium	119 (34)	96 (39)	105 (35)
High	121 (35)	85 (34)	107 (36)
Missing	12 (3)	4 (2)	12 (4)
Duration of epilepsy, years			
Mean (SD)	16.1 (14.3)	16.9 (15.7)	16.2 (14.4)
Median (IQR)	12 (5–22)	12 (5–22)	11 (5–23)
Missing, *n* (%)	46 (13)	29 (12)	41 (14)
Number of seizure last years, *n* (%)			
No seizure	235 (68)	165 (67)	213 (71)
Seizures (1 or above)	96 (28)	69 (28)	74 (25)
Missing	15 (4)	13 (5)	13 (4)
Side effects			
Mean (SD)	1.56 (0.79)	1.45 (0.67)	1.53 (0.76)
Median (IQR)	1 (1–2)	1 (1–2)	1 (1–2)
Missing, *n* (%)	7 (2)	2 (1)	7 (2)
Well-being (WHO-5)			
Mean (SD)	68.9 (18.9)	68.0 (19.4)	69.0 (18.9)
Median (IQR)	72 (60–80)	72 (56–80)	72 (60–80)
Missing, *n* (%)	10 (3)	3 (1)	9 (3)
General health			
Mean (SD)	2.62 (0.92)	2.64 (0.91)	2.59 (0.92)
Median (IQR)	3 (2–3)	3 (2–3)	3 (2–3)
Missing, *n* (%)	7 (2)	3 (1)	7 (2)
Social support for health (HLQ, subscale 4)			
Mean (SD)	3.30 (0.54)	3.32 (0.56)	3.30 (0.55)
Median (IQR)	3.4 (3–3.8)	3.4 (3–3.8)	3.4 (3–3.8)
Missing, *n* (%)	15 (4)	6 (2)	14 (5)
Ability to actively engage with healthcare providers (HLQ, subscale 6)			
Mean (SD)	3.87 (0.84)	3.82 (0.91)	3.87 (0.84)
Median (IQR)	4 (3.4–4.5)	4 (3.4–4.6)	4 (3.4–4.4)
Missing, *n* (%)	18 (5)	6 (2)	16 (5)
Understanding health information well enough to know what to do (HLQ, subscale 9)			
Mean (SD)	4.01 (0.79)	3.94 (0.86)	4.02 (0.77)
Median (IQR)	4 (3.6–4.6)	4 (3.6–4.6)	4 (3.6–4.6)
Missing, *n* (%)	18 (5)	6 (2)	16 (5)

Table 1 Participant baseline characteristics, *N* = 593 *(Continued)*

Variables	Intention-to-treat population		Per-protocol population
	Intervention arm (open access telePRO): *n* = 346	Control arm (standard telePRO): *n* = 247	Patients who completed the open access intervention: *n* = 300
Self-efficacy (GSE)			
Mean (SD)	29.35 (6.08)	29.23 (6.73)	29.71 (5.86)
Missing, *n* (%)	16 (5)	7 (3)	15 (5)

Differences between the arms were evaluated by chi-squared test, Wilcoxon Mann–Whitney test, or unpaired *t*-test. No *p*-value < 0.05 was found
SD Standard deviation, *IQR* Interquartile range, *WHO-5* WHO-5 well-being index, *HLQ* Health literacy questionnaire, *GSE* General self-efficacy scale

(95% CI: – 1.29 to – 0.04) and – 0.21 (95% CI: – 0.38 to – 0.03), respectively, compared to the standard arm. However, the mental well-being was lower in the open access arm than in the standard arm, difference: – 5.95 (95% CI: – 10.81 to – 1.08). No statistically significant differences were found in the high age group.

Attrition and sensitivity analyses

Web (*N* = 648) and paper responders (*N* = 395) of the baseline research questionnaire were compared (Fig. 1). Web-responders were younger, had a higher level of education, and higher scores of health literacy and self-efficacy. No differences were found in gender, well-being, and general health. Responders (*N* = 352) to the follow-up research questionnaire were compared to non-responders (*N* = 241) by using data gathered at baseline. Participants who did not respond were younger, mean (SD) age 43.2 (17.0) years versus 49.0 (17.0) years, *P* = 0.0001. Furthermore, non-responders had a lower WHO-5 well-being score, mean (SD) 65.6 (19.1) versus 70.5 (18.8), *P* = 0.003, lower scores of the HLQ 6 "Ability to actively engage with healthcare providers", mean (SD) 3.75 (0.90) versus 3.91 (0.83), *P* = 0.04, lower self-reported general health, *P* = 0.04, and lower level of education, *P* = 0.02. No differences were found with respect to gender, cohabitation status, "social support for health" (HLQ 4), "understand health information well enough to know what to do" (HLQ 9), self-efficacy (GSE),

Table 2 Healthcare utilization during an 18-month follow-up period among outpatients with epilepsy

Primary outcome	Intention-to-treat population			Per-protocol population	
	Intervention arm (open access telePRO): *N* = 343	Control arm (standard telePRO): *N* = 243	Mean difference (95% CI)	Completed the open access intervention: *N* = 300	Mean difference (95% CI)
Outpatient visits [a]					
Mean (SD)	0.45 (0.95)	0.42 (0.86)	0.03 (–0.11 to 0.18)	0.43 (0.91)	0.04 (–0.12 to 0.20)
Median (Range)	0 (0–7)	0 (0–6)		0 (0–7)	
Telephone consultations [a]					
Mean (SD)	0.99 (1.88)	1.30 (2.46)	–0.32 (– 0.68 to 0.05)	0.90 (1.80)	–0.20 (– 0.55 to 0.15)
Median (Range)	0 (0–12)	1 (0–22)		0 (0–12)	
Hospitalizations [a]					
Mean (SD)	0.05 (0.29)	0.09 (0.49)	–0.04 (– 0.10 to 0.03)	0.05 (0.25)	0.0002 (–0.05 to 0.05)
Median (Range)	0 (0–3)	0 (0–5)		0 (0–2)	
Emergency room visits [b]					
Mean (SD)	0.07 (0.38)	0.19 (0.72)	–0.11 (– 0.21 to –0.01)	0.06 (0.31)	–0.08 (– 0.18 to 0.007)
Median (Range)	0 (0–4)	0 (0–7)		0 (0–3)	

The estimated intention-to-treat mean differences and 95% CIs were obtained after simple linear regression by using the bootstrap method with 1000 replications [39]
The estimated per-protocol mean differences and 95% CIs were obtained after multiple linear regression adjusted for gender, age, education, cohabitation status, epilepsy duration, and seizures last year by using the bootstrap method with 1000 replications [39]
SD Standard deviation, *CI* Confidence interval
[a]at the Department of Neurology, Aarhus University Hospital, [b] at Aarhus University Hospital

Table 3 Patient-reported outcomes measured 18 months after randomization among outpatients with epilepsy

Secondary outcomes	Intention-to-treat population			Per-protocol population	
	Intervention arm (open access telePRO): *N* = 202	Control arm (standard telePRO): *N* = 150	Difference [a] at 18- mo. follow-up (95% CI)	Completed the open access intervention: *N* = 195	Difference [b] at 18- mo. follow-up (95% CI)
Well-being (WHO-5)					
Mean (SD)	66.99 (19.45)	69.29 (18.01)	−3.21 (−6.38 to −0.05)	66.94 (19.64)	−3.26 (−6.68 to 0.16)
Missing, *n* (%)	4 (2)	4 (3)		3 (2)	
Social support for health (HLQ 4)					
Mean (SD)	3.24 (0.60)	3.38 (0.53)	−0.08 (−0.17 to 0.02)	3.24 (0.61)	−0.04(− 0.14 to 0.07)
Missing, *n* (%)	5 (2)	4 (3)		5 (3)	
Ability to actively engage with healthcare providers (HLQ 6)					
Mean (SD)	3.84 (0.82)	3.87 (0.89)	−0.05 (−0.21 to 0.10)	3.85 (0.82)	−0.04 (− 0.20 to 0.14)
Missing, *n* (%)	6 (3)	4 (3)		6 (3)	
Understanding health information well enough to know what to do (HLQ 9)					
Mean (SD)	4.03 (0.77)	3.97 (0.85)	0.009 (−0.13 to 0.15)	4.02 (0.78)	0.04 (−0.12 to 0.20)
Missing, *n* (%)	6 (3)	4 (3)		6 (3)	
Self-efficacy (GSE)					
Mean (SD)	29.78 (5.69)	29.73 (6.14)	−0.22 (−1.22 to 0.78)	29.81 (5.75)	−0.02 (−1.16 to 1.13)
Missing, *n* (%)	7 (3)	4 (3)		7 (4)	
General health					
Mean (SD)	2.63 (0.93)	2.60 (0.82)	0.05 (−0.10 to 0.19)	2.61 (0.94)	0.06 (−0.11 to 0.22)
Missing, *n* (%)	1 (0.05)	1 (0.07)		1 (0.05)	
No. of seizure last year					
Mean (SD)	2.50 (11.89)	3.20 (10.21)	−0.72 (−3.20 to 1.75)	2.52 (12.04)	−0.63 (−3.50 to 2.24)
Missing, *n* (%)	36 (18)	28 (19)		33 (17)	
Side effects					
Mean (SD)	1.54 (0.76)	1.56 (0.83)	−0.03 (−0.18 to 0.11)	1.53 (0.77)	0.005 (−0–15 to 0.17)
Missing, *n* (%)	6 (3)	1 (0.07)		6 (3)	
Patient activation [c]					
Mean (SD)	3.42 (0.65)	3.34 (0.77)	0.04 (−0.10 to 0.17)	3.42 (0.65)	0.001 (−0.15 to 0.15)
Missing, *n* (%)	6 (3)	4 (3)		6 (3)	
Patient activation [d]					
Mean (SD)	3.22 (0.72)	3.12 (0.75)	0.01 (−0.13 to 0.16)	3.22 (0.73)	0.02 (−0.14 to 0.17)
Missing, *n* (%)	5 (2)	4 (3)		5 (2)	
Confidence					
Mean (SD)	1.39 (0.65)	1.33 (0.53)	0.03 (−0.9 to 0.16)	1.39 (0.65)	0.06 (−0.07 to 0.20)
Missing, *n* (%)	21 (10)	9 (6)		20 (10)	
Safety					
Mean (SD)	1.41 (0.70)	1.35 (0.56)	0.02 (−0.12 to 0.16)	1.41 (0.70)	0.07 (−0.09 to 0.23)
Missing, *n* (%)	37 (18)	14 (9)		36 (18)	
Satisfaction					
Mean (SD)	1.63 (0.68)	1.61 (0.59)	0.01 (−0.13 to 0.15)	1.63 (0.69)	0.05 (−0.11 to 0.20)
Missing, *n* (%)	35 (17)	18 (12)		34 (17)	

SD Standard deviation, *CI* Confidence interval, *WHO-5* WHO-5 well-being index, *HLQ* Health literacy questionnaire, *GSE* General self-efficacy scale

[a]The estimated intention-to-treat differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure

[b]The estimated per-protocol differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure, gender, age, education, cohabitation status, epilepsy duration, and seizures last year

[c]I am confident that I can tell when I need to get outpatient care

[d]I am confident I can figure out solutions when new situations or problems arise with my health condition

side effects, seizures during last year, and duration of epilepsy.

The sensitivity analyses showed that the difference in the WHO-5 score changed to not statistically significant if the missing data at follow-up were based on a 5-point lower baseline value in the standard arm, but were unchanged in the open access arm, 0.11 (95% CI: − 1.84 to 2.07) and a 5-point higher baseline value in the open access arm, but were unchanged in the standard arm, 0.18 (95% CI: − 1.81 to 2.17) (Additional file 3, page 10). The difference became stronger if the missing data was based on a 5-point higher baseline value in the standard arm, but were unchanged in the open access arm, − 3.83 (95% CI: − 5.78 to − 1.88) and a 5-point lower baseline value in the open access arm, but were unchanged in the standard arm, − 4.02 (95% CI: − 5.94 to − 2.10).

Process evaluation in the open access arm

Overall, activity in terms of number of logins to the "My Epilepsy" web site and questionnaires filled in initiated by the patients decreased during the follow-up period (Fig. 2). At the same time, an increased number of reminders were sent to patients; the response rate (37%) was, however, low.

Discussion

The open access telePRO intervention showed no statistically significant differences in use of health care resources in terms of telephone consultations, outpatient visits, or hospital admissions compared to standard telePRO follow-up. The open access arm had a statistically significant, slightly lower mean number of emergency room visits. No statistically significant differences were found in clinical outcome measures such as mortality, number of seizures, and side effects. Further, for patient-centered outcome measures, the mental well-being was statistically significantly lower in the open access arm than in the standard arm, but there were no differences in general health status. No statistically significant differences were found in patient self-management and health service evaluation.

Effectiveness in terms of use of health care resources in an open access or a telemedicine intervention has been investigated in other studies, which have reported a lower number of outpatient visits in the intervention group than in the control group in patients with inflammatory bowel disease [37, 40, 41] and rheumatoid arthritis [42, 43]. However, in these studies the control group was offered pre-scheduled follow-up appointments in the clinic. This was not the case in our study, as the standard arm was another model of PRO-based follow-up, not traditional follow-up with fixed appointments. As a consequence, the contrast between the two arms was relatively low in our study, which may explain why there was small or no differences between the two types of PRO-based follow-up. Preferably, both study arms in this study should be compared with patients who received face-to-face follow-up. However, since standard telePRO was standard care for 2500 outpatients with epilepsy at Aarhus University Hospital in 2016, this option was limited. By including patients already attending standard telePRO follow-up, we took advantage of recruiting a large proportion of prevalent patients within 6 months. Although, another study could be based on new patients, as they could be randomized to the two different PRO-based follow-up arms used in this study or traditional face-to-face follow-up. However, this design was not customized according to clinical practice in the department and recruitment of new patients would require a much longer recruitment period, as the department receives approximately only 250 new patients yearly. Still, we found it relevant to investigate whether a more tailored individual follow-up method initiated by the patients could lead to further benefits for the patients and the health care system, as the fixed-interval PRO-based model is primary driven by clinicians who determine the questionnaire intervals. Our findings related to clinical outcomes are in accordance with results from other studies in patients with rheumatoid arthritis [42, 43], inflammatory bowel disease [41], and chronic obstructive pulmonary disease [44], which also found no differences in disease activity. In addition, a review reported similar health-related quality of life and psychological outcomes in patient-initiated follow-up compared to traditional follow-up [18]. Patients in the open access arm in our study reported a statistically lower self-reported mental well-being than those in the standard arm; however, the difference was small and probably not clinically significant, since a clinically relevant change on the WHO-5 scale is considered to be 10 points [27]. Further, the measurement error of WHO-5 has been estimated to be around 20 points in an epilepsy outpatient population, and this should be taken into consideration when using the scale to measure change over time [45].

The main strength of this study was the pragmatic randomized design customized according to real life implementation of PRO-based follow-up in clinical practice. Furthermore, the study included a large study sample, and loss to follow-up in the analyses of the primary outcome was limited. However, some limitations should be noted. The baseline level of the HLQ subscales and the GSE in the epilepsy population was nearly the same as the in Danish population as a whole [31, 32]; thus a ceiling effect occurred, and it became difficult to observe improvement in these constructs over time. Another limitation was the low response rate of the questionnaires at 18-month follow-up. Only approximately 60% of the patients in both arms responded; thus, selection bias cannot be ruled out. The sensitivity analyses of self-reported WHO-5 well-being showed that the results could potentially be both underestimated or moved toward the null hypothesis of no effect. All

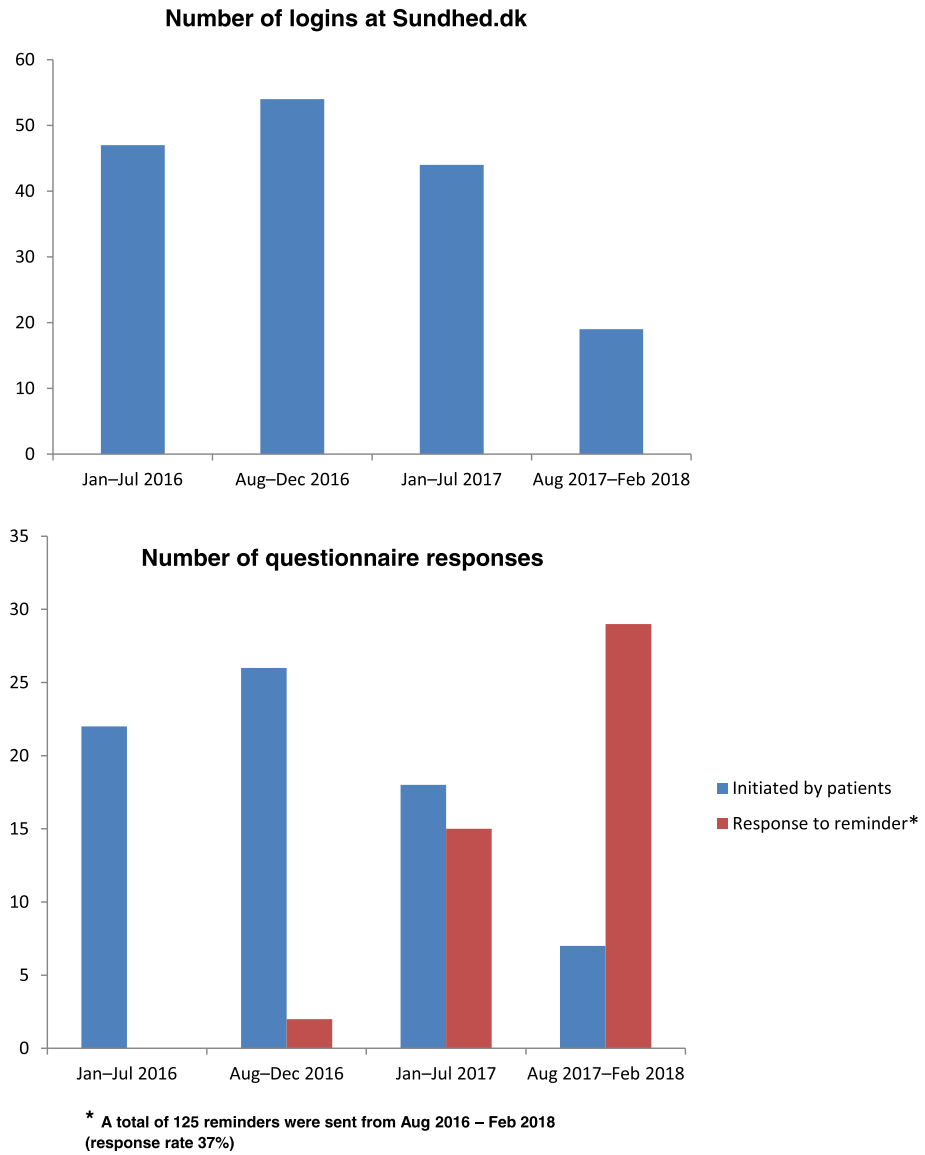


Fig. 2 Activity in the open access arm ($N = 346$) during follow-up in terms of logins to the "My Epilepsy" web site at Sundhed.dk, questionnaires responses initiated by patients, and questionnaire responses of reminders sent to patients

eligible patients were referred to standard telePRO; therefore, we decided to use a pre-randomization design with few inclusion criteria. However, this study only included patients who were able to fill in the questionnaire via the Internet. As shown in Fig. 1, this is a selected standard telePRO group of patients because 395 paper responders (38%) were excluded. This should be taken into consideration in the generalization of the results.

Non-adherence to the use of health technology interventions is a common problem [46]. Participants make use of the intervention differently and not all continue to use the intervention as intended [46]. Data from Fig. 2 indicate that the number of patients who actually used the open access website was low. It is important to explain program failures if the intervention does not function as intended. Greenhalgh et al. describe the complexities of predicting if or how people engage with health technology in a new theoretical framework [47]. The framework includes seven domains: the condition, the technology, the value proposition, the adopter system, the organization, the wider societal context, and the interaction between these domains over time [47]. We have used elements of this framework to discuss challenges related to the open access intervention. At the organizational level, much effort was put into developing the intervention, but the implementation strategy was probably insufficient, leading to issues related to knowledge and confidence in using the intervention as intended by the patients. Information

about the intervention was mailed to the participants only once during the follow-up period, and the participants were expected to take action by themselves if they declined to participate in the open access arm. At the individual patient level, the open access intervention demanded some self-management skills because the patients were expected to actively interact with the health care system. However, patients are only activated if "they understand their role in the health care process and have the knowledge, skills and confidence to carry it out" [2]. There could be resistance by the patients regarding filling in a questionnaire in order to get in contact with the clinic, as they might have found it easier to call the clinic if they needed to talk to a clinician. Resistance could also be related to technical issues, for example, the login procedure to the open access website required some extra steps, and technical problems were experienced by some patients. Qualitative data regarding the patient perspective will be further investigated.

Conclusion

There is growing need for health care strategies to manage more effective and patient-centered care. This study did not find, as hypothesized, less use of health care resources or improved patient self-management or satisfaction in the patient-initiated PRO-based initiative compared to fixed-interval PRO-based follow-up. Patient-initiated PRO-based follow-up may be used as an alternative to fixed-interval PRO-based follow-up in epilepsy outpatients who prefer this approach, but there is insufficient evidence for recommending a system-wide shift to patient-initiated PRO-based follow-up. How patients are allocated to this health care service is important, and individuals' self-management skills should be taken into consideration. Further work should explore the effects of using a patient-initiated PRO-based intervention in clinical practice, preferably, comparing these patients with patients using a fixed-appointment follow-up procedure.

Additional files

Additional file 1: Information for Reporting Randomized Controlled Trials With Patient report Outcomes. (PDF 1129 kb)

Additional file 2: Disease-specific epilepsy questionnaire. (PDF 1501 kb)

Additional file 3: Stratified analyses, sensitivity analyses, and the randomization computer code. (PDF 624 kb)

Abbreviations

CI: Confidence interval; CONSORT: Consolidated Standard of Reporting Trial; GH: General health; GSE: General Self-Efficacy Scale; HLQ: Health Literacy Questionnaire; IQR: Interquartile range; ITT: Intention-to-treat; PAM: Patient Activation Measure; PRO: Patient-reported outcome; SD: Standard deviation; WHO-5: WHO-5 Well-Being Index

Acknowledgements

We would like to thank all the clinical staff and the patients at the epilepsy outpatient clinic at Aarhus University Hospital.

Authors' contributions

Study design: LMVS, CTM, KL, HTM, AT and NHH. Data collection and analyses: LMVS and NHH. All authors were involved in the manuscript preparation and all authors read and approved the final manuscript.

Funding

This work was supported by Aarhus University, the Central Denmark Regions Health Research Foundation, and TrygFonden. These funding souces were not involved in the design, data collection or analyses, interpretation of data or writing the manuscript.

Availability of data and materials

Access to the protocol for this study can be found as reference [22] and via this link: https://doi.org/10.1186/s12913-017-2015-8. Please contact the corresponding author for further guidance regarding data request.

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Danish Data Protection Agency (reference number 1–16–02-691-14). The Ethics Committee of Central Denmark Region was contacted and decided that approval was not a requirement for the study. All data were stored and treated with confidentiality. Data were stored in the WestChronic/AmbuFlex web-system, which is physically located in the Central Denmark Region's Server Park protected behind the firewall and Treat Management Gateway [10].

Consent for publication

Not applicable.

Competing interests

JC received honoraria serving on the scientific advisory board of UCB Nordic and Eisai AB, received honoraria giving lectures for UCB Nordic and Eisai AB, and received funding for a trip from UCB Nordic. The other authors have no competing interests.

Author details

[1]AmbuFlex/WestChronic, Occupational Medicine, University Research Clinic, Aarhus University, Gl. Landevej 61, DK-7400 Herning, Denmark. [2]Department of Neurology, Aarhus University Hospital, Palle Juul-Jensens Boulevard 165, DK-8200 Aarhus N, Denmark. [3]Department of Clinical Medicine, Aarhus University, Palle Juul-Jensens Boulevard 82, DK-8200 Aarhus N, Denmark. [4]Department of Public Health, Aarhus University, Bartholins Allé 2, DK-8000 Aarhus, Denmark. [5]Department of Rheumatology, Aarhus University Hospital, Palle Juul-Jensens Boulevard 59, DK-8200 Aarhus N, Denmark. [6]Department of Clinical Epidemiology, Aarhus University Hospital, Olof Palmes Allé 43-45, DK-8200 Aarhus N, Denmark.

Received: 16 January 2019 Accepted: 27 August 2019
Published online: 13 September 2019

References

1. Barlow, J., Wright, C., Sheasby, J., Turner, A., & Hainsworth, J. (2002). Self-management approaches for people with chronic conditions: A review. *Patient Education and Counseling, 48*, 177–187.
2. Santana, M. J., & Feeny, D. (2014). Framework to assess the effects of using patient-reported outcome measures in chronic care management. *Quality of Life Research, 23*, 1505–1513.
3. US Department of Health and Human Services Food and Drug Administration (2009) Guidance for industry: patient-reported outcome measures: Use in medical product development to support labeling claims. Available via www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf Accessed 23 May 2018.
4. Valderas, J. M., Kotzeva, A., Espallargues, M., Guyatt, G., Ferrans, C. E., Halyard, M. Y., Revicki, D. A., Symonds, T., Parada, A., & Alonso, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: A systematic review of the literature. *Quality of Life Research, 17*, 179–193.

5. Marshall, S., Haywood, K., & Fitzpatrick, R. (2006). Impact of patient-reported outcome measures on routine practice: A structured review. *Journal of Evaluation in Clinical Practice, 12*, 559–568.
6. Chen, J., Ou, L., & Hollis, S. J. (2013). A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Services Research, 13*, 211. https://doi.org/10.1186/1472-6963-13-211.
7. Johansen, M. A., Berntsen, G. K., Schuster, T., Henriksen, E., & Horsch, A. (2012). Electronic symptom reporting between patient and provider for improved health care service quality: A systematic review of randomized controlled trials. Part 2: Methodological quality and effects. *Journal of Medical Internet Research, 14*, e126.
8. Schougaard, L. M., Larsen, L. P., Jessen, A., Sidenius, P., Dorflinger, L., de Thurah, A., & Hjollund, N. H. (2016). AmbuFlex: Tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. *Quality of Life Research, 25*, 525–534.
9. Mejdahl, C. T., Schougaard, L. M. V., Hjollund, N. H., Riiskjaer, E., Thorne, S., & Lomborg, K. (2018). PRO-based follow-up as a means of self-management support - An interpretive description of the patient perspective. *Journal of Patient-Reported Outcomes, 2*, 38. https://doi.org/10.1186/s41687-018-0067-0 eCollection 2017.
10. Hjollund, N. H., Larsen, L. P., Biering, K., Johnsen, S. P., Riiskjaer, E., & Schougaard, L. M. (2014). Use of patient-reported outcome (PRO) measures at group and patient levels: Experiences from the generic integrated PRO system, WestChronic. *Interactive Journal of Medical Research, 3*, e5.
11. Fisher, R. S., van Emde Boas, W., Blume, W., Elger, C., Genton, P., Lee, P., & Engel, J., Jr. (2005). Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia, 46*, 470–472.
12. Josephson, C. B., Patten, S. B., Bulloch, A., Williams, J. V. A., Lavorato, D., Fiest, K. M., Secco, M., & Jette, N. (2017). The impact of seizures on epilepsy outcomes: A national, community-based survey. *Epilepsia, 58*, 764–771.
13. Jacoby, A., & Baker, G. A. (2008). Quality-of-life trajectories in epilepsy: A review of the literature. *Epilepsy and Behavior, 12*, 557–571.
14. Nixon, A., Kerr, C., Breheny, K., & Wild, D. (2013). Patient Reported Outcome (PRO) assessment in epilepsy: A review of epilepsy-specific PROs according to the Food and Drug Administration (FDA) regulatory requirements. *Health and Quality of Life Outcomes, 11*, 38. https://doi.org/10.1186/1477-7525-11-38.
15. Moura, L. M., Schwamm, E., Moura Junior, V., Seitz, M. P., Hsu, J., Cole, A. J., & Schwamm, L. H. (2016). Feasibility of the collection of patient-reported outcomes in an ambulatory neurology clinic. *Neurology, 87*, 2435–2442.
16. Bergmann, M., Prieschl, M., Walser, G., Luef, G., Rumpold, G., & Unterberger, I. (2018). Computer-based monitoring and evaluation of epilepsy-related health variables and their impact on treatment decision. *Epilepsy and Behavior, 84*, 173–178.
17. Whear, R., Abdul-Rahman, A. K., Boddy, K., Thompson-Coon, J., Perry, M., & Stein, K. (2013). The clinical effectiveness of patient initiated clinics for patients with chronic or recurrent conditions managed in secondary care: A systematic review. *PLoS One, 8*, e74774.
18. Whear, R., Abdul-Rahman, A. K., Thompson-Coon, J., Boddy, K., Perry, M. G., & Stein, K. (2013). Patient initiated clinics for patients with chronic or recurrent conditions managed in secondary care: A systematic review of patient reported outcomes and patient and clinician satisfaction. *BMC Health Services Research, 13*, 501. https://doi.org/10.1186/1472-6963-13-501.
19. Taneja, A., Su'a, B., & Hill, A. G. (2014). Efficacy of patient-initiated follow-up clinics in secondary care: A systematic review. *Internal Medicine Journal, 44*, 1156–1160.
20. Schulz, K. F., Altman, D. G., Moher, D., & CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology, 63*, 834–840.
21. Calvert, M., Blazeby, J., Altman, D. G., Revicki, D. A., Moher, D., Brundage, M. D., & CONSORT PRO Group. (2013). Reporting of patient-reported outcomes in randomized trials: The CONSORT PRO extension. *JAMA, 309*, 814–822.
22. Schougaard, L. M., Mejdahl, C. T., Petersen, K. H., Jessen, A., de Thurah, A., Sidenius, P., Lomborg, K., & Hjollund, N. H. (2017). Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: Study protocol for a pragmatic randomised controlled study. *BMC Health Services Research, 17*, 83. https://doi.org/10.1186/s12913-017-2015-8.
23. Zelen, M. (1979). A new design for randomized clinical trials. *The New England Journal of Medicine, 300*, 1242–1245.
24. Dumville, J. C., Hahn, S., Miles, J. N., & Torgerson, D. J. (2006). The use of unequal randomisation ratios in clinical trials: A review. *Contemporary Clinical Trials, 27*, 1–12.
25. Schougaard, L. M. V., de Thurah, A., Christiansen, D. H., Sidenius, P., & Hjollund, N. H. (2018). Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: A test-retest reliability study. *BMJ Open, 8*, e021337. https://doi.org/10.1136/bmjopen-2017-021337.
26. Staehr Johansen, K. (1998). *The use of well-being measures in primary health care – the Dep-Care project; in World Health Organization, regional office for Europe: Well-being measures in primary health care – the DepCare project*. Geneva: World Health Organization Target 12, E60246.
27. Topp, C. W., Ostergaard, S. D., Sondergaard, S., & Bech, P. (2015). The WHO-5 well-being index: A systematic review of the literature. *Psychotherapy and Psychosomatics, 84*, 167–176.
28. Bjorner, J. B., Damsgaard, M. T., Watt, T., & Groenvold, M. (1998). Tests of data quality, scaling assumptions, and reliability of the Danish SF-36. *Journal of Clinical Epidemiology, 51*, 1001–1011.
29. Bjorner, J. B., Thunedborg, K., Kristensen, T. S., Modvig, J., & Bech, P. (1998). The Danish SF-36 health survey: Translation and preliminary validity studies. *Journal of Clinical Epidemiology, 51*, 991–999.
30. Osborne, R. H., Batterham, R. W., Elsworth, G. R., Hawkins, M., & Buchbinder, R. (2013). The grounded psychometric development and initial validation of the health literacy questionnaire (HLQ). *BMC Public Health, 13*, 658. https://doi.org/10.1186/1471-2458-13-658.
31. Maindal, H. T., Kayser, L., Norgaard, O., Bo, A., Elsworth, G. R., & Osborne, R. H. (2016). Cultural adaptation and validation of the health literacy questionnaire (HLQ): Robust nine-dimension Danish language confirmatory factor model. *Springerplus, 5*, 1232. https://doi.org/10.1186/s40064-016-2887-9 eCollection 2016.
32. Scholz, U., Gutiérrez-Doña, B., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment, 18*(3), 242–251.
33. Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). Windsor: NFER-NELSON.
34. Maindal, H. T., Sokolowski, I., & Vedsted, P. (2009). Translation, adaptation and validation of the American short form patient activation measure (PAM13) in a Danish version. *BMC Public Health, 9*, 209. https://doi.org/10.1186/1471-2458-9-209.
35. Danish Cancer Society (2017) Spørgeskema. En undersøgelse af dine behov og oplevelser under udredning og behandling for kræft (In Danish). Available via https://www.cancer.dk/dyn/resources/File/file/2/6912/1513860458/final_barometerskema_nyeste.pdf (Accessed Nov 2018).
36. Frank, L. (2000). Epidemiology. When an entire country is a cohort. *Science, 287*, 2398–2399.
37. Williams, J. G., Cheung, W. Y., Russell, I. T., Cohen, D. R., Longo, M., & Lervy, B. (2000). Open access follow up for inflammatory bowel disease: Pragmatic randomised trial and cost effectiveness study. *BMJ, 320*, 544–548.
38. Hesterberg, T. C. (2015). What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum. *The American Statistician, 69*, 371–386.
39. Stata Corporation, College Station, Texas. Bootstrap manual. Available via https://www.stata.com/manuals13/rbootstrap.pdf Accessed 27 June 2019.
40. Robinson, A., Thompson, D. G., Wilkin, D., Roberts, C., & Northwest Gastrointestinal Research Group. (2001). Guided self-management and patient-directed follow-up of ulcerative colitis: A randomised trial. *The Lancet, 358*, 976–981.
41. de Jong, M. J., van der Meulen-de Jong, A. E., Romberg-Camps, M. J., Becx, M. C., Maljaars, J. P., Cilissen, M., van Bodegraven, A. A., Mahmmod, N., Markus, T., Hameeteman, W. M., Dijkstra, G., Masclee, A. A., Boonen, A., Winkens, B., van Tubergen, A., Jonkers, D. M., & Pierik, M. J. (2017). Telemedicine for management of inflammatory bowel disease (myIBDcoach): A pragmatic, multicentre, randomised controlled trial. *The Lancet, 390*, 959–968.
42. de Thurah, A., Stengaard-Pedersen, K., Axelsen, M., Fredberg, U., Schougaard, L. M. V., Hjollund, N. H. I., Pfeiffer-Jensen, M., Laurberg, T. B., Tarp, U., Lomborg, K., & Maribo, T. (2018). Tele-health Followup strategy for tight control of disease activity in rheumatoid arthritis: Results of a randomized controlled trial. *Arthritis Care and Research (Hoboken), 70*, 353–360.

43. Hewlett, S., Kirwan, J., Pollock, J., Mitchell, K., Hehir, M., Blair, P. S., Memel, D., & Perry, M. G. (2005). Patient initiated outpatient follow up in rheumatoid arthritis: Six year randomised controlled trial. *BMJ, 330*, 171.
44. Berkhof, F. F., Hesselink, A. M., Vaessen, D. L., Uil, S. M., Kerstjens, H. A., & van den Berg, J. W. (2014). The effect of an outpatient care on-demand-system on health status and costs in patients with COPD. A randomized trial. *Respiratory Medicine, 108*, 1163–1170.
45. Schougaard, L. M. V., de Thurah, A., Bech, P., Hjollund, N. H., & Christiansen, D. H. (2018). Test-retest reliability and measurement error of the Danish WHO-5 well-being index in outpatients with epilepsy. *Health and Quality of Life Outcomes, 16*, 175. https://doi.org/10.1186/s12955-018-1001-0.
46. Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. (2012). Persuasive system design does matter: A systematic review of adherence to web-based interventions. *Journal of Medical Internet Research, 14*, e152.
47. Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research, 19*, e367.

Publisher's Note

Table 1. Information for Reporting Randomized Controlled Trials With Patient reported Outcomes

Section/Topic	Item	CONSORT 2010 Statement Checklist Item	PRO-Specific Extensions Are Prefaced by the letter *P*	
		Title and Abstract		
	1a	Identification as a randomized trial in the title		Page 1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[3]	P1b: The PRO should be identified in the abstract as a primary or secondary outcome	Page 1-2
		Introduction		
Background and objectives	2a	Scientific background and explanation of rationale	Including background and rationale for PRO assessment	Page 3-4
	2b	Specific objectives or hypotheses	P2b: The PRO hypothesis should be stated and relevant domains identified, if applicable	Page 4-5
		Methods		
Trial design	3a	Description of trial design (such as parallel, factorial), including allocation ratio		Page 5
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		N/A
Participants	4a	Eligibility criteria for participants	Not PRO-specific, unless the PROs were used in eligibility or stratification criteria	Page 5
	4b	Settings and locations where the data were collected		Page 6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered		Page 6-8
Outcomes	6a	Completely defined prespecified primary and secondary outcome measures, including how and when they were assessed	P6a: Evidence of PRO instrument validity and reliability should be provided or cited if available including the person completing the PRO and methods of data collection (paper, telephone, electronic, other)	Page 8-9
	6b	Any changes to trial outcomes after the trial commenced, with reasons		N/A
Sample size	7a	How sample size was determined	Not required for PRO unless it is a primary study outcome	Page 10
	7b	When applicable, explanation of any interim analyses and stopping guidelines		N/A
		Randomization		
Sequence generation	8a	Method used to generate the random allocation sequence		Page 6
	8b	Type of randomization; details of any restriction (such as blocking and block size)		Page 6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned		Page 5-6
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions		Page 6
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		Page 6
	11b	If relevant, description of the similarity of interventions		N/A
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	P12a: Statistical approaches for dealing with missing data are explicitly stated	Page 10-11
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		Page 11
		Results		
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome	The number of PRO outcome data at baseline and at subsequent time points should be made transparent	Figure 1
	13b	For each group, losses and exclusions after randomization, together with reasons		Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up		Page 12
	14b	Why the trial ended or was stopped		N/A
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Including baseline PRO data when collected	Table 1
Numbers analyzed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Required for PRO results	Figure 1 + Table 2 and 3
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, the estimated effect size, and its precision (such as 95% confidence interval)	For multidimensional PRO results from each domain and time point	Page 12-13 + Table 2 and 3
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		N/A
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory	Including PRO analyses, where relevant	Page 13-14
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)		N/A
		Discussion		
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	P20/21: PRO–specific limitations and implications for generalizability and clinical practice	Page 15-17
Generalizability	21	Generalizability (external validity, applicability) of the trial findings		Page 16
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	PRO data should be interpreted in relation to clinical outcomes including survival data, where relevant	Page 15-17
		Other Information		
Registration	23	Registration number and name of trial registry		In Abstract
Protocol	24	Where the full trial protocol can be accessed, if available		Page 18
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		Page 18

Additional file 3

Supplemental analyses

Analyses stratified on age

AGE HIGH (Age at baseline ≥ 45.7 years)

Health care utilization during an 18-month follow-up period among outpatients with epilepsy

Primary outcomes	**Open access (ITT)** ***N*** **= 169**	**Control** ***N*** **= 124**	**ITT Mean difference (95%CI)**
Outpatient visits [a]			
Mean (SD)	0.41 (0.97)	0.38 (0.88)	0.04 (−0.18 to 0.25)
Median (Range)	0 (0–7)	0 (0–6)	
Telephone consultations [a]			
Mean (SD)	0.86 (1.62)	0.85 (1.39)	0.01 (−0.33 to 0.35)
Median (Range)	0 (0–10)	1 (0–10)	
Hospitalizations [a]			
Mean (SD)	0.08 (0.36)	0.06 (0.26)	0.02 (−0.05 to 0.09)
Median (Range)	0 (0–3)	0 (0–2)	
Emergency room visits [b]			
Mean (SD)	0.08 (0.41)	0.10 (0.42)	−0.02 (−0.12 to 0.08)
Median (Range)	0 (0–4)	0 (0–3)	

[a] at the Department of Neurology, Aarhus University Hospital, [b] at Aarhus University Hospital
SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat
The estimated IIT mean differences and 95% CIs were obtained after simple linear regression by using the bootstrap method with 1000 replications.

AGE LOW (Age at baseline < 45.7 years)

Health care utilization during an 18-month follow-up period among outpatients with epilepsy

Primary outcomes	**Open access (ITT)** ***N*** **= 174**	**Control** ***N*** **= 119**	**ITT Mean difference (95%CI)**
Outpatient visits [a]			
Mean (SD)	0.48 (0.93)	0.45 (0.84)	0.03 (−0.18 to 0.24)
Median (Range)	0 (0–6)	0 (0–5)	
Telephone consultations [a]			
Mean (SD)	1.11 (2.09)	1.78 (3.16)	−0.67 (−1.29 to −0.04)
Median (Range)	0 (0–12)	1 (0–22)	
Hospitalizations [a]			
Mean (SD)	0.02 (0.18)	0.12 (0.65)	−0.09 (−0.22 to 0.03)
Median (Range)	0 (0–2)	0 (0–5)	
Emergency room visits [b]			
Mean (SD)	0.06 (0.34)	0.27 (0.94)	−0.21 (−0.38 to −0.03)
Median (Range)	0 (0–3)	0 (0–7)	

[a] at the Department of Neurology, Aarhus University Hospital, [b] at Aarhus University Hospital
SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat
The estimated IIT mean differences and 95% CIs were obtained after simple linear regression by using the bootstrap method with 1000 replications.

Additional file 3

Analyses stratified on age

AGE HIGH (Age at baseline ≥ 45.7 years)

Patient-reported outcomes measured 18 months after randomization among outpatients with epilepsy

Secondary outcomes	*N*	Open access (ITT)	*N*	Control	ITT: Difference [a] at 18-mo. follow-up (95%CI)
Well-being (WHO-5) Mean (SD)	107	69.83(19.46)	89	69.48(17.69)	−1.36 (−5.59 to 2.87)
Self-efficacy (GSE) Mean (SD)	105	30.00(5.52)	90	29.56(5.95)	−0.03 (−1.33 to 1.27)
HLQ 4 Mean (SD)	106	3.24(0.61)	90	3.35(0.53)	−0.04 (−0.17 to 0.09)
HLQ 6 Mean (SD)	105	3.89(0.79)	91	3.85(0.90)	−0.07 (−0.27 to 0.14)
HLQ 9 Mean (SD)	105	4.01(0.79)	90	3.90(0.88)	−0.007 (−0.19 to 0.17)
General health Mean (SD)	109	2.68(0.88)	92	2.66(0.73)	0.04 (−0.14 to 0.23)
No. of seizure last year Mean (SD)	93	2.71(13.00)	76	3.58(11.00)	−0.69 (−4.02 to 2.67)
Treatment side effects Mean (SD)	108	1.53(0.77)	92	1.55(0.89)	−0.05(−0.26 to 0.15)
Patient activation [b] Mean (SD)	108	3.39(0.61)	91	3.27(0.79)	0.08 (−0.09 to 0.25)
Patient activation [c] Mean (SD)	108	3.23(0.68)	91	3.10(0.80)	0.03(−0.15 to 0.22)
Confidence Mean (SD)	97	1.38(0.64)	87	1.30(0.51)	0.01(−0.14 to 0.17)
Safety Mean (SD)	88	1.43(0.67)	82	1.28(0.48)	0.10(−0.07 to 0.27)
Satisfaction Mean (SD)	91	1.65(0.69)	79	1.59(0.59)	0.05 (−0.14 to 0.24)

[a] The estimated intention-to-treat differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure

[b] I am confident that I can tell when I need to get outpatient care

[c] I am confident I can figure out solutions when new situations or problems arise with my health condition

SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat; WHO-5: WHO-5 Well-being Index; GSE: General Self-efficacy Scale; HLQ: Health Literacy Questionnaire

Additional file 3

Analyses stratified on age

AGE LOW (Age at baseline < 45.7 years)

Patient-reported outcomes measured 18 months after randomization among outpatients with epilepsy

Secondary outcomes	*N*	Open access (ITT)	*N*	Control	ITT: Difference [a] at 18-mo. follow-up (95%CI)
Well-being (WHO-5) Mean (SD)	91	63.65(19.01)	57	68.98(18.66)	−5.95 (−10.81 to −1.08)
Self-efficacy (GSE) Mean (SD)	90	29.52(5.90)	56	30.02(6.47)	−0.50 (−2.11 to 1.12)
HLQ 4 Mean (SD)	88	3.25(0.61)	53	3.42(0.54)	−0.14 (−0.29 to 0.006)
HLQ 6 Mean (SD)	88	3.80(0.82)	55	3.90(0.87)	−0.02 (−0.26 to 0.21)
HLQ 9 Mean (SD)	88	4.04(0.76)	55	4.07(0.82)	0.005 (−0.22 to 0.23)
General health Mean (SD)	92	2.57(0.99)	57	2.51(0.95)	0.05 (−0.19 to 0.30)
No. of seizure last year Mean (SD)	73	2.23(10.39)	46	2.52(8.83)	−0.58 (−4.27 to 3.12)
Treatment side effects Mean (SD)	88	1.55(0.76)	57	1.56(0.71)	−0.002(−0.20 to 0.19)
Patient activation [b] Mean (SD)	88	3.45(0.69)	55	3.45(0.74)	−0.06 (−0.29 to 0.17)
Patient activation [c] Mean (SD)	89	3.20(0.77)	55	3.16 (0.66)	−0.03(−0.27 to 0.20)
Confidence Mean (SD)	84	1.40(0.66)	54	1.37(0.56)	0.05(−0.15 to 0.25)
Safety Mean (SD)	77	1.39(0.73)	54	1.46(0.66)	−0.09(−0.33 to 0.15)
Satisfaction Mean (SD)	76	1.61(0.67)	53	1.62(0.60)	−0.03 (−0.24 to 0.18)

[a] The estimated intention-to-treat differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure

[b] I am confident that I can tell when I need to get outpatient care

[c] I am confident I can figure out solutions when new situations or problems arise with my health condition

SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat; WHO-5: WHO-5 Well-being Index; GSE: General Self-efficacy Scale; HLQ: Health Literacy Questionnaire

Additional file 3

Analyses stratified on gender

FEMALE

Health care utilization during an 18-month follow-up period among outpatients with epilepsy

Primary outcomes	Open access (ITT) *N* = 162	Control *N* = 130	ITT Mean difference (95%CI)
Outpatient visits [a]			
Mean (SD)	0.55 (1.01)	0.52 (1.04)	0.03 (−0.20 to 0.26)
Median (Range)	0 (0–7)	0 (0–6)	
Telephone consultations [a]			
Mean (SD)	1.27 (2.17)	1.62 (2.93)	−0.36 (−0.95 to 0.24)
Median (Range)	0 (0–12)	1 (0–22)	
Hospitalizations [a]			
Mean (SD)	0.02 (0.19)	0.11 (0.56)	−0.08 (−0.18 to 0.01)
Median (Range)	0 (0–2)	0 (0–5)	
Emergency room visits [b]			
Mean (SD)	0.04 (0.22)	0.19 (0.75)	−0.16 (−0.29 to −0.02)
Median (Range)	0 (0–2)	0 (0–7)	

[a] at the Department of Neurology, Aarhus University Hospital, [b] at Aarhus University Hospital
SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat
The estimated IIT mean differences and 95% CIs were obtained after simple linear regression by using the bootstrap method with 1000 replications.

MALE

Health care utilization during an 18-month follow-up period among outpatients with epilepsy

Primary outcomes	Open access (ITT) *N* = 181	Control *N* = 113	ITT Mean difference (95%CI)
Outpatient visits [a]			
Mean (SD)	0.36 (0.88)	0.29 (0.58)	0.07 (−0.10 to 0.23)
Median (Range)	0 (0–6)	0 (0–2)	
Telephone consultations [a]			
Mean (SD)	0.74 (1.54)	0.94 (1.71)	−0.20 (−0.59 to 0.19)
Median (Range)	0 (0–10)	1 (0–10)	
Hospitalizations [a]			
Mean (SD)	0.07 (0.35)	0.06 (0.41)	0.01 (−0.08 to 0.10)
Median (Range)	0 (0–3)	0 (0–4)	
Emergency room visits [b]			
Mean (SD)	0.10 (0.48)	0.18 (0.70)	−0.07 (−0.22 to 0.07)
Median (Range)	0 (0–4)	0 (0–6)	

[a] at the Department of Neurology, Aarhus University Hospital, [b] at Aarhus University Hospital
SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat
The estimated IIT mean differences and 95% CIs were obtained after simple linear regression by using the bootstrap method with 1000 replications.

Additional file 3

Analyses stratified on gender

FEMALE

Patient-reported outcomes measured 18 months after randomization among outpatients with epilepsy

Secondary outcomes	*N*	Open access (ITT)	*N*	Control	ITT: Difference [a] at 18-mo. follow-up (95%CI)
Well-being (WHO-5) Mean (SD)	91	66.24(18.96)	79	67.24(18.91)	−2.94 (−7.32 to 1.44)
Self-efficacy (GSE) Mean (SD)	87	29.29(5.86)	80	29.21(6.19)	−0.46 (−1.81 to 0.89)
HLQ 4 Mean (SD)	89	3.31(0.58)	78	3.35(0.59)	−0.009 (−0.14 to 0.12)
HLQ 6 Mean (SD)	87	3.89(0.77)	79	3.93(0.79)	−0.11 (−0.29 to 0.07)
HLQ 9 Mean (SD)	87	4.06(0.81)	79	4.03(0.76)	−0.005 (−0.18 to 0.17)
General health Mean (SD)	93	2.59(0.89)	80	2.70(0.83)	−0.05 (−0.26 to 0.16)
No. of seizure last year Mean (SD)	78	2.32(9.60)	65	3.60(10.80)	−1.51 (−4.97 to 1.95)
Treatment side effects Mean (SD)	90	1.51(0.71)	80	1.59(0.85)	−0.14 (−0.32 to 0.05)
Patient activation [b] Mean (SD)	90	3.46(0.66)	79	3.39(0.72)	0.08 (−0.12 to 0.27)
Patient activation [c] Mean (SD)	91	3.26(0.73)	79	3.08(0.73)	0.12 (−0.08 to 0.32)
Confidence Mean (SD)	86	1.34(0.59)	76	1.29(0.51)	−0.02(−0.18 to 0.13)
Safety Mean (SD)	80	1.34(0.57)	74	1.36(0.56)	−0.08(−0.25 to 0.10)
Satisfaction Mean (SD)	82	1.65(0.64)	69	1.58(0.55)	0.04 (−0.15 to 0.24)

[a] The estimated intention-to-treat differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure

[b] I am confident that I can tell when I need to get outpatient care

[c] I am confident I can figure out solutions when new situations or problems arise with my health condition

SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat; WHO-5: WHO-5 Well-being Index; GSE: General Self-efficacy Scale; HLQ: Health Literacy Questionnaire

Additional file 3

Analyses stratified on gender

MALE

Patient-reported outcomes measured 18 months after randomization among outpatients with epilepsy

Secondary outcomes	*N*	Open access (ITT)	*N*	Control	ITT: Difference [a] at 18-mo. follow-up (95%CI)
Well-being (WHO-5) Mean (SD)	107	67.63(19.93)	67	71.70(16.70)	−3.74 (−8.42 to 0.93)
Self-efficacy (GSE) Mean (SD)	90	30.18(5.55)	66	30.36(6.06)	−0.09 (−1.59 to 1.42)
HLQ 4 Mean (SD)	105	3.19(0.62)	65	3.41(0.50)	−0.15 (−0.29 to −0.002)
HLQ 6 Mean (SD)	106	3.82(0.83)	67	3.79(0.99)	0.05 (−0.19 to 0.30)
HLQ 9 Mean (SD)	106	3.99(0.75)	66	3.89(0.96)	0.04 (−0.18 to 0.27)
General health Mean (SD)	108	2.66(0.97)	69	2.49(0.80)	0.12 (−0.09 to 0.34)
No. of seizure last year Mean (SD)	88	2.66(13.66)	57	2.70(9.57)	0.24 (−2.27 to 2.75)
Treatment side effects Mean (SD)	106	1.56(0.81)	69	1.52(0.80)	0.05 (−0.17 to 0.28)
Patient activation [b] Mean (SD)	106	3.39(0.64)	67	3.28(0.83)	−0.007 (−0.21 to 0.19)
Patient activation [c] Mean (SD)	106	3.18(0.74)	67	3.18 (0.78)	−0.10(−0.32 to 0.11)
Confidence Mean (SD)	95	1.44(0.70)	65	1.37(0.55)	0.08 (−0.12 to 0.27)
Safety Mean (SD)	85	1.48(0.80)	62	1.34(0.57)	0.12(−0.10 to 0.34)
Satisfaction Mean (SD)	85	1.61(0.73)	63	1.63(0.63)	−0.02 (−0.23 to 0.18)

[a] The estimated intention-to-treat differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure

[b] I am confident that I can tell when I need to get outpatient care

[c] I am confident I can figure out solutions when new situations or problems arise with my health condition

SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat; WHO-5: WHO-5 Well-being Index; GSE: General Self-efficacy Scale; HLQ: Health Literacy Questionnaire

Additional file 3

Analyses stratified on health literacy 'Social support for health" (HLQ4)

HLQ 4 HIGH (HLQ 4 scale measured at baseline ≥ 3.4)

Health care utilization during an 18-month follow-up period among outpatients with epilepsy

Primary outcomes	**Open access (ITT)** *N* **= 183**	**Control** *N* **= 136**	**ITT Mean difference (95%CI)**
Outpatient visits [a]			
Mean (SD)	0.43 (0.91)	0.40 (0.93)	0.03 (−0.16 to 0.23)
Median (Range)	0 (0–6)	0 (0–6)	
Telephone consultations [a]			
Mean (SD)	1.04 (1.78)	1.31 (2.18)	−0.27 (−0.72 to 0.19)
Median (Range)	0 (0–10)	1 (0–12)	
Hospitalizations [a]			
Mean (SD)	0.03 (0.21)	0.05 (0.49)	−0.02 (−0.07 to 0.03)
Median (Range)	0 (0–2)	0 (0–2)	
Emergency room visits [b]			
Mean (SD)	0.05 (0.24)	0.15 (0.45)	−0.11 (−0.19 to −0.02)
Median (Range)	0 (0 – 2)	0 (0 – 3)	

[a] at the Department of Neurology, Aarhus University Hospital, [b] at Aarhus University Hospital
SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat
The estimated IIT mean differences and 95% CIs were obtained after simple linear regression by using the bootstrap method with 1000 replications.

HLQ 4 LOW (HLQ 4 scale measured at baseline < 3.4)

Health care utilization during an 18-month follow-up period among outpatients with epilepsy

Primary outcomes	**Open access (ITT)** *N* **= 175**	**Control** *N* **= 113**	**ITT Mean difference (95%CI)**
Outpatient visits [a]			
Mean (SD)	0.45 (0.96)	0.42 (0.75)	0.03 (−0.17 to 0.23)
Median (Range)	0 (0–7)	0 (0–3)	
Telephone consultations [a]			
Mean (SD)	0.92 (1.92)	1.32 (2.76)	−0.40 (−0.98 to 0.18)
Median (Range)	0 (0–12)	1 (0–22)	
Hospitalizations [a]			
Mean (SD)	0.06 (0.34)	0.12 (0.67)	−0.06 (−0.19 to0.07)
Median (Range)	0 (0–3)	0 (0–5)	
Emergency room visits [b]			
Mean (SD)	0.10 (0.48)	0.22 (0.94)	−0.12 (−0.31 to 0.06)
Median (Range)	0 (0–4)	0 (0–7)	

[a] at the Department of Neurology, Aarhus University Hospital, [b] at Aarhus University Hospital
SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat
The estimated IIT mean differences and 95% CIs were obtained after simple linear regression by using the bootstrap method with 1000 replications.

Additional file 3

Analyses stratified on health literacy 'Social support for health" (HLQ4)

HLQ 4 HIGH (HLQ 4 scale measured at baseline ≥ 3.4)

Patient-reported outcomes measured 18 months after randomization among outpatients with epilepsy

Secondary outcomes	N	Open access (ITT)	N	Control	ITT: Difference [a] at 18-mo. follow-up (95%CI)
Well-being (WHO-5) Mean (SD)	115	70.64(17.39)	89	69.29(18.01)	−3.68 (−7.39 to 0.03)
Self-efficacy (GSE) Mean (SD)	112	31.01(5.72)	90	30.96(5.42)	−0.03 (−1.32 to 1.26)
General health Mean (SD)	115	2.42(0.87)	92	2.41(0.71)	0.06 (−0.12 to 0.25)
No. of seizure last year Mean (SD)	98	2.19(11.36)	78	2.69(9.84)	−0.34 (−3.60 to 2.92)
Treatment side effects Mean (SD)	115	1.46(0.72)	92	1.52(0.84)	−0.05(−0.23 to 0.13)
Patient activation [b] Mean (SD)	113	3.54(0.64)	90	3.46(0.75)	0.06 (−0.12 to 0.23)
Patient activation [c] Mean (SD)	113	3.33(0.70)	90	3.21(0.79)	−0.003(−0.20 to 0.19)
Confidence Mean (SD)	104	1.31(0.58)	86	1.29(0.51)	−0.03(−0.17 to 0.11)
Safety Mean (SD)	98	1.34(0.67)	82	1.30(0.51)	0.02(−0.16 to 0.19)
Satisfaction Mean (SD)	98	1.51(0.63)	81	1.56(0.63)	−0.03 (−0.21 to 0.15)

[a] The estimated intention-to-treat differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure

[b] I am confident that I can tell when I need to get outpatient care

[c] I am confident I can figure out solutions when new situations or problems arise with my health condition

SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat; WHO-5: WHO-5 Well-being Index; GSE: General Self-efficacy Scale; HLQ: Health Literacy Questionnaire

Additional file 3

Analyses stratified on health literacy 'Social support for health" (HLQ4)

HLQ 4 LOW (HLQ 4 scale measured at baseline < 3.4)

Patient-reported outcomes measured 18 months after randomization among outpatients with epilepsy

Secondary outcomes	N	Open access (ITT)	N	Control	ITT: Difference [a] at 18-mo. follow-up (95%CI)
Well-being (WHO-5) Mean (SD)	93	63.31(20.62)	61	64.98(20.63)	−2.73 (−8.08 to 2.62)
Self-efficacy (GSE) Mean (SD)	93	28.32(5.04)	59	27.76(6.60)	−0.26 (−1.83 to 1.31)
General health Mean (SD)	96	2.85(0.93)	61	2.87(0.92)	0.03 (−0.20 to 0.26)
No. of seizure last year Mean (SD)	81	2.60(11.99)	47	3.83(10.58)	−2.04 (−5.43 to 1.35)
Treatment side effects Mean (SD)	91	1.60(0.80)	61	1.61(0.78)	−0.02(−0.25 to 0.21)
Patient activation [b] Mean (SD)	93	3.28(0.61)	58	3.17(0.78)	0.06 (−0.16 to 0.29)
Patient activation [c] Mean (SD)	94	3.11(0.71)	58	3.00 (0.68)	0.03(−0.20 to 0.26)
Confidence Mean (SD)	85	1.51(0.72)	57	1.39(0.56)	0.10(−0.11 to 0.32)
Safety Mean (SD)	73	1.49(0.71)	56	1.43(0.63)	0.01(−0.22 to 0.25)
Satisfaction Mean (SD)	77	1.75(0.71)	53	1.68(0.51)	0.05 (−0.17 to 0.28)

[a] The estimated intention-to-treat differences and 95% CIs were obtained after multiple linear regression adjusted for baseline measure

[b] I am confident that I can tell when I need to get outpatient care

[c] I am confident I can figure out solutions when new situations or problems arise with my health condition

SD: Standard deviation; CI: Confidence interval; ITT: Intention-to-treat; WHO-5: WHO-5 Well-being Index; GSE: General Self-efficacy Scale; HLQ: Health Literacy Questionnaire

Additional file 3

Sensitivity analyses of WHO-5 Well-being Index

Sensitivity analyses were performed to establish the impact of missing self-reported data in WHO-5 Well-Being Index. The response rate of the follow-up questionnaire was approximately 60%.

Mean and standard deviation (SD) of the WHO-5 Well-being Index measured at baseline and at follow-up (18 months after randomization):

	Intervention arm (open access telePRO) *N* = 346	**Control arm (standard telePRO) *N* = 247**
Baseline WHO-5 Mean (SD) Missing, *n (%)*	 68.9 (18.9) 10 (3)	 68.0 (19.4) 3 (1)
Follow-up WHO-5 Mean (SD) Missing, *n (%)*	 67.0 (19.45) 144 (42)	 69.3 (18.01) 97 (39)

If the WHO-5 score was missing at follow-up, the score was imputed by using the WHO-5 score from the baseline questionnaire. Four scenarios regarding the imputed follow-up values were considered:

Analysis 1
Intervention arm: Baseline WHO reduced with 5 points
Control arm: WHO-5 baseline unchanged

Analysis 2
Intervention arm: WHO-5 baseline unchanged
Control arm: Baseline WHO reduced with 5 points

Analysis 3
Intervention arm: Baseline WHO increased with 5 points
Control arm: WHO-5 baseline unchanged

Analysis 4
Intervention arm: WHO-5 baseline unchanged
Control arm: Baseline WHO increased with 5 points

Then, between-arm differences in the WHO-5 score at follow-up were analyzed by multiple linear regression adjusted for the baseline WHO-5 value. Patients who died or emigrated ($N = 7$) or had missing WHO-5 scores at both baseline and follow-up ($N=3$) were not included in the analyses.

WHO-5 Well-being	**Intervention arm *N* = 341**	**Control arm *N* = 242**	**Difference at 18 mo. follow-up (95% Confidence interval)**
Analysis 1 WHO-5 mean (SD)	 64.4 (19.5)	 67.7 (18.6)	 −4.02 (−5.94 to −2.10)
Analysis 2 WHO-5 mean (SD)	 66.5 (19.3)	 65.8 (19.0)	 0.11 (−1.84 to 2.07)
Analysis 3 WHO-5 mean (SD)	 68.5 (19.3)	 67.7 (18.6)	 0.18 (−1.81 to 2.17)
Analysis 4 WHO-5 mean (SD)	 66.5 (19.3)	 69.7 (18.5)	 −3.83 (−5.78 to −1.88)

Additional file 3

The randomization computer code

The randomization was programmed in the computer language PHP:

```
$fraction= round(mt_rand() / mt_getrandmax(),4);  //generates random decimal fraction between 0.000 and 1.000

IF($fraction>0.45)

$allocation='OpenAccess';

ELSE

$allocation='Normal';
```

APPENDICES

Appendix 1: Detailed search strategy within themes and papers

Appendix 2: The AmbuFlex/epilepsy questionnaire

Appendix 3: Patient information in Study IV [In Danish]

Appendix 4: Co-authorship declarations

APPENDIX 1

Detailed search strategy within themes and papers

APPENDIX 1: Detailed search strategy within themes and papers

Theme 1: Effect of using PRO measures in clinical practice

PRO	Clinical practice
Patient Reported Outcome Measures (Mesh)	Outpatient clinics, Hospital (Mesh)
Patient reported outcome* (Text Word)	Ambulatory Care (Mesh)
Self reported health status (Text Word)	Clinical Decision-Making (Mesh)
Self reported outcome* (Text Word)	Monitoring, Ambulatory (Mesh)
Self reported patient outcome* (Text Word)	Secondary Care (Mesh)
Selfreported (Text Word)	Primary Health Care (Mesh)
ePROM (Text Word)	Clinical practice (Text Word)
ePRO (Text Word)	Follow up (Text Word)
Electronic symptom reporting (Text Word)	Clinical care (Text Word)
Symptom reporting (Text Word)	Clinical setting (Text Word)
Patient reported data (Text Word)	Oncologic setting
	Health care service* (Text Word)
	Feedback (Text Word)

((((((((((("Patient Reported Outcome Measures"[Mesh]) OR patient reported outcome*[Text Word]) OR Self reported health status[Text Word]) OR Self reported outcome*[Text Word]) OR Self reported patient outcome*[Text Word]) OR Selfreported[Text Word]) OR ePROM[Text Word]) OR ePRO[Text Word]) OR symptom reporting[Text Word]) OR Electronic symptom reporting[Text Word]) OR Patient reported data[Text Word])) AND ((((((((((("Outpatient Clinics, Hospital"[Mesh]) OR "Ambulatory Care"[Mesh]) OR "Clinical Decision-Making"[Mesh]) OR "Monitoring, Ambulatory"[Mesh]) OR "Secondary Care"[Mesh]) OR "Primary Health Care"[Mesh]) OR clinical practice[Text Word]) OR follow up[Text Word]) OR clinical care[Text Word]) OR clinical setting[Text Word]) OR oncologic setting[Text Word]) OR health care service*[Text Word]) OR feedback[Text Word])

Theme 2: Effect of using PRO measures in remote patient management

PRO	Remote monitoring
Patient Reported Outcome Measures (Mesh)	Telemedicine (Mesh)
Patient reported outcome* (Text Word)	Remote monitoring (Text Word)
Self reported health status (Text Word)	Remote patient monitoring (Text Word)
Self reported outcome* (Text Word)	Remote patient management (Text Word)
Self reported patient outcome* (Text Word)	Remote patient follow-up (Text Word)
Selfreported (Text Word)	Remote follow-up (Text Word)
ePROM (Text Word)	Symptom monitoring (Text Word)
ePRO (Text Word)	Mediated follow up (Text Word)
Electronic symptom reporting (Text Word)	Ambulatory Care (Mesh)
Symptom reporting (Text Word)	Outpatient clinic, Hospital (Mesh)
Patient reported data (Text Word)	Monitoring, Ambulatory (Mesh)

((((((((((("Patient Reported Outcome Measures"[Mesh]) OR patient reported outcome*[Text Word]) OR Self reported health status[Text Word]) OR Self reported outcome*[Text Word]) OR Self reported patient outcome*[Text Word]) OR Selfreported[Text Word]) OR ePROM[Text Word]) OR ePRO[Text Word]) OR symptom reporting[Text Word]) OR Electronic symptom reporting[Text Word]) OR Patient reported data[Text Word])) AND ((((((((((("Telemedicine"[Mesh]) OR remote monitoring[Text Word]) OR Remote patient monitoring[Text Word]) OR Remote patient management[Text Word]) OR Remote patient follow-up[Text Word]) OR Remote follow-up[Text Word]) OR Symptom monitoring[Text Word]) OR Mediated follow up[Text Word]) OR "Ambulatory Care"[Mesh]) OR "Outpatient Clinics, Hospital"[Mesh]) OR "Monitoring, Ambulatory"[Mesh])

APPENDIX 1: Detailed search strategy within themes and papers

Theme 3 + Paper IV: Effect of using PRO measures in patient-initiated follow-up

Patient-initiated	PRO
Patient initiated follow up (Text Word)	Patient Reported Outcome Measures (Mesh)
Patient initiated followup (Text Word)	Patient reported outcome* (Text Word)
Patient initiated clinics (Text Word)	Self reported health status (Text Word)
Open access follow up (Text Word)	Self reported outcome* (Text Word)
Open access followup (Text Word)	Self reported patient outcome* (Text Word)
Direct access follow up (Text Word)	Selfreported (Text Word)
Direct access followup (Text Word)	ePROM (Text Word)
	ePRO (Text Word)
	Electronic symptom reporting (Text Word)
	Symptom reporting (Text Word)
	Patient reported data (Text Word)

((((((Patient initiated follow up[Text Word]) OR Patient initiated followup[Text Word]) OR Patient initiated clinics[Text Word]) OR Open access follow up[Text Word]) OR Open access followup[Text Word]) OR Direct access follow up[Text Word]) OR Direct access followup[Text Word]

(((((((((Patient initiated follow up[Text Word]) OR Patient initiated followup[Text Word]) OR Patient initiated clinics[Text Word]) OR Open access follow up[Text Word]) OR Open access followup[Text Word]) OR Direct access follow up[Text Word]) OR Direct access followup[Text Word])) AND ((((((((((("Patient Reported Outcome Measures"[Mesh]) OR patient reported outcome*[Text Word]) OR Self reported health status[Text Word]) OR Self reported outcome*[Text Word]) OR Self reported patient outcome*[Text Word]) OR Selfreported[Text Word]) OR ePROM[Text Word]) OR ePRO[Text Word]) OR symptom reporting[Text Word]) OR Electronic symptom reporting[Text Word]) OR Patient reported data[Text Word]))

Paper I: Questionnaire non-response characteristics

PRO & Questionnaire	Non-response
Patient Reported Outcome Measures (Mesh)	Non respon* (Text Word)
Patient reported outcome* (Text Word)	
Self reported health status (Text Word)	
Self reported outcome* (Text Word)	
Self reported patient outcome* (Text Word)	
Selfreported (Text Word)	
ePROM (Text Word)	
ePRO (Text Word)	
Electronic symptom reporting (Text Word)	
Symptom reporting (Text Word)	
Patient reported data (Text Word)	
Surveys and questionnaires (Mesh)	

((((((((((((("Patient Reported Outcome Measures"[Mesh]) OR patient reported outcome*[Text Word]) OR Self reported health status[Text Word]) OR Self reported outcome*[Text Word]) OR Self reported patient outcome*[Text Word]) OR Selfreported[Text Word]) OR ePROM[Text Word]) OR ePRO[Text Word]) OR symptom reporting[Text Word]) OR Electronic symptom reporting[Text Word]) OR Patient reported data[Text Word]) OR ("Surveys and Questionnaires"[Mesh]))) AND Non respon*[Text Word])

APPENDIX 1: Detailed search strategy within themes and papers

Paper II: Test-retest reliability of PRO-algorithm used in clinical decision support

PRO	Clinical practice	Epilepsy
Patient Reported Outcome Measures (Mesh)	Outpatient clinics, Hospital (Mesh)	Epilepsy (Mesh)
Patient reported outcome* (Text Word)	Ambulatory Care (Mesh)	
Self reported health status (Text Word)	Clinical Decision-Making (Mesh)	
Self reported outcome* (Text Word)	Monitoring, Ambulatory (Mesh)	
Self reported patient outcome* (Text Word)	Secondary Care (Mesh)	
Selfreported (Text Word)	Primary Health Care (Mesh)	
ePROM (Text Word)	Clinical practice (Text Word)	
ePRO (Text Word)	Follow up (Text Word)	
Electronic symptom reporting (Text Word)	Clinical care (Text Word)	
Symptom reporting (Text Word)	Clinical setting (Text Word)	
Patient reported data (Text Word)	Oncologic setting	
	Health care service* (Text Word)	
	Feedback (Text Word)	

((((((((((("Patient Reported Outcome Measures"[Mesh]) OR patient reported outcome*[Text Word]) OR Self reported health status[Text Word]) OR Self reported outcome*[Text Word]) OR Self reported patient outcome*[Text Word]) OR Selfreported[Text Word]) OR ePROM[Text Word]) OR ePRO[Text Word]) OR symptom reporting[Text Word]) OR Electronic symptom reporting[Text Word]) OR Patient reported data[Text Word])) AND "Epilepsy"[Mesh])

((((((((((("Patient Reported Outcome Measures"[Mesh]) OR patient reported outcome*[Text Word]) OR Self reported health status[Text Word]) OR Self reported outcome*[Text Word]) OR Self reported patient outcome*[Text Word]) OR Selfreported[Text Word]) OR ePROM[Text Word]) OR ePRO[Text Word]) OR symptom reporting[Text Word]) OR Electronic symptom reporting[Text Word]) OR Patient reported data[Text Word])) AND ((((((((((("Outpatient Clinics, Hospital"[Mesh]) OR "Ambulatory Care"[Mesh]) OR "Clinical Decision-Making"[Mesh]) OR "Monitoring, Ambulatory"[Mesh]) OR "Secondary Care"[Mesh]) OR "Primary Health Care"[Mesh]) OR clinical practice[Text Word]) OR follow up[Text Word]) OR clinical care[Text Word]) OR clinical setting[Text Word]) OR oncologic setting[Text Word]) OR health care service*[Text Word]) OR feedback[Text Word])) AND "Epilepsy"[Mesh]

APPENDIX 1: Detailed search strategy within themes and papers

Paper III: Test-retest reliability of WHO-5 Well-Being Index

WHO-5 Well-Being Index	Reliability	Epilepsy
WHO-5 well-being index	Validation studies as topic [Mesh]	Epilepsy (Mesh)
WHO-5 well-being index [Text Word]	Reproducibility of results [Mesh]	
WHO-5 well-being	Reliability [Text Word]	
WHO-5 well-being [Text Word]	Reproducibility [Text Word]	
WHO-5 wellbeing	Test-retest relilability [Text Word]	
WHO-5 wellbeing [Text Word]	Test-retest study [Text Word]	
WHO-5 index	Test-retest [Text Word]	
WHO-5 index [Text Word]	Measurement error [Text Word]	
WHO-5 score*		
WHO-5 score* [Text Word]		
WHO-5 scale*		
WHO-5 scale* [Text Word]		
WHO-5 questionnaire		
WHO-5 questionnaire [Text Word]		
WHO-5		
WHO-5 [Text Word]		

((((((((((((((((WHO-5 well-being index) OR WHO-5 well-being index[Text Word]) OR WHO-5 well-being) OR WHO-5 well-being[Text Word]) OR WHO-5 wellbeing) OR WHO-5 wellbeing[Text Word]) OR WHO-5 index) OR WHO-5 index[Text Word]) OR WHO-5 score*) AND WHO-5 score*[Text Word]) OR WHO-5 scale*) OR WHO-5 scale*[Text Word]) OR WHO-5 questionnaire) OR WHO-5 questionnaire[Text Word]) OR WHO-5) OR WHO-5[Text Word])) AND ((((((((((("Validation Studies as Topic"[Mesh]) OR "Reproducibility of Results"[Mesh]) OR Reliability[Text Word]) OR Reproducibility[Text Word]) OR test retest reliability) OR Test-retest study) OR Test-retest study[Text Word]) OR Test-retest) OR Test-retest[Text Word]) OR Measurement error) OR Measurement error[Text Word])

((((((((((((((((((WHO-5 well-being index) OR WHO-5 well-being index[Text Word]) OR WHO-5 well-being) OR WHO-5 well-being[Text Word]) OR WHO-5 wellbeing) OR WHO-5 wellbeing[Text Word]) OR WHO-5 index) OR WHO-5 index[Text Word]) OR WHO-5 score*) AND WHO-5 score*[Text Word]) OR WHO-5 scale*) OR WHO-5 scale*[Text Word]) OR WHO-5 questionnaire) OR WHO-5 questionnaire[Text Word]) OR WHO-5) OR WHO-5[Text Word])) AND ((((((((((("Validation Studies as Topic"[Mesh]) OR "Reproducibility of Results"[Mesh]) OR Reliability[Text Word]) OR Reproducibility[Text Word]) OR test retest reliability) OR Test-retest study) OR Test-retest study[Text Word]) OR Test-retest) OR Test-retest[Text Word]) OR Measurement error) OR Measurement error[Text Word]))) AND "Epilepsy"[Mesh]

((((((((((((((((WHO-5 well-being index) OR WHO-5 well-being index[Text Word]) OR WHO-5 well-being) OR WHO-5 well-being[Text Word]) OR WHO-5 wellbeing) OR WHO-5 wellbeing[Text Word]) OR WHO-5 index) OR WHO-5 index[Text Word]) OR WHO-5 score*) AND WHO-5 score*[Text Word]) OR WHO-5 scale*) OR WHO-5 scale*[Text Word]) OR WHO-5 questionnaire) OR WHO-5 questionnaire[Text Word]) OR WHO-5) OR WHO-5[Text Word])) AND "Epilepsy"[Mesh]

APPENDIX 2

The AmbuFlex/epilepsy questionnaire

When did you have your most recent epileptic seizure? (Write year and month - write "?" if you do not remember)

Year (f. ex. 2011):: ☐☐☐☐

Month no. (f. ex. 11): ☐☐

How many seizures did you have last year? (if none, write 0)

Number of attacks: : ☐☐

How many absence seizures have you had in the last 3 months?

Number of attacks: : ☐☐

How many generalized seizures (convulsions) have you had during the last 3 months?

Number of attacks: : ☐☐

Answer the following questions only if you have had at least 1 epileptic seizure during the last year

Do your epileptic seizures occur during sleep?	☐ Yes	☐ No	
Are your seizures getting worse?	☐ Yes	☐ No	
Have you sustained an injury during a seizure?	☐ No	☐ Yes, but not a serious one	☐ Serious damage (f. ex. bone fracture, cut wounds)
Have you been in contact with an emergency room because of epilepsy since your last visit to the outpatient department?	☐ Yes	☐ No	

Are your relatives worried about you because of your epilepsy?

Put one tick	☐ Never	☐ Rarely	☐ Occasionally	☐ Frequently	☐ Do not know	☐ Not applicable

During the last 4 weeks to what degree have you suffered from:

Headache	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Dizziness	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Tremor/shacking	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Double vision or other visual disturbances	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Loss of appetite	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Eating too much	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Difficulty remembering	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Difficulty concentrating	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
A feeling that you easily become aggressive	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ very often
Severe fatigue	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Sadness	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Being afraid of having a new seizure during the next weeks	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Lack of interest or pleasure in sexual activity	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often
Have you in the last 4 weeks had suicidal thoughts?	☐ Never	☐ Occasionally	☐ Sometimes	☐ Often	☐ Very often

Please indicate for each of the five statements which is closest to how you have been feeling over the last two weeks.

Over the last two weeks

I have felt cheerful and in good spirits	☐ All of the time	☐ Most of the time	☐ More than half of the time	☐ Less than half of the time	☐ Some of the time	☐ At no time
I have felt calm and relaxed	☐ All of the time	☐ Most of the time	☐ More than half of the time	☐ Less than half of the time	☐ Some of the time	☐ At no time
I have felt active and vigorous	☐ All of the time	☐ Most of the time	☐ More than half of the time	☐ Less than half of the time	☐ Some of the time	☐ At no time
I woke up feeling fresh and rested	☐ All of the time	☐ Most of the time	☐ More than half of the time	☐ Less than half of the time	☐ Some of the time	☐ At no time
My daily life has been filled with things that interest me	☐ All of the time	☐ Most of the time	☐ More than half of the time	☐ Less than half of the time	☐ Some of the time	☐ At no time

In general, would you say your health is:

Put one tick	☐ Excellent	☐ Very good	☐ Good	☐ Fair	☐ Poor

Compared to one year ago, how would you rate your health in general now?

Put one tick	☐ Much better now than one year ago	☐ Somewhat better now than one year ago	☐ About the same	☐ Somewhat worse now than one year ago	☐ Much worse than one year ago

Do you have other diseases or conditions that have a greater effect on your health than your epilepsy?

☐ Yes

☐ No

The next questions deal with your medical treatment

How often do you think you have forgotten to take some of your medicine?	☐ Daily	☐ Weekly	☐ Monthly	☐ Very rarely, never
Does your epilepsy medicine have side effects?	☐ No	☐ Yes, a few	☐ Yes, some	☐ Yes, many

The next questions deal with work (being a student counts as work)

Have you felt stressed at your work in the last 12 months?	☐ No	☐ Yes, a bit	☐ Yes, a lot	☐ I'm no longer in the work force	☐ I'm out of work
How much do you now work compared with 12 months ago?	☐ I work more	☐ About the same	☐ I work less	☐ I'm no longer in the work force	☐ I'm out of work
If you work less now, is it because of your epilepsy?	☐ Yes	☐ Partly	☐ No		

Has your epilepsy put serious limitations on your life?

☐ No

☐ Yes, describe how: _____

How much alcohol do you drink on average in the course of a week? (Refers to beer, wine and spirits. If you drink less than 1 unit a week, write 0)

Write number of units: ____

Do you use recreational drugs? (f. ex. hash)

Put one tick	☐ Never	☐ Monthly	☐ Weekly	☐ Daily

The next questions are only relevant for women

Are you pregnant?

☐ Yes

☐ No

Do you plan to get pregnant within the next 12 months?

☐ Yes

☐ No

How much do you weight? (number of kg without clothes)

Write answer here: ___

Have you driven a car in the last month?

- ☐ Yes
- ☐ No

What is your present need for contact with the outpatient clinic?

- ☐ I phone myself if I need to talk to someone
- ☐ I'd rather have someone phone me
- ☐ I'd like to have an outpatient appointment
- ☐ I don't know

Who has filled in this questionnaire?

- ☐ I have filled in the questionnaire
- ☐ I have had help filling in the questionnaire
- ☐ Someone else has filled in the questionnaire for me (f. ex., spouse, contact person)

May we phone you regarding your answers to the questionnaire?

- ☐ No
- ☐ Yes - and my phone number is: _____

Here you can write a short note to the personnel that read the questionnaire

APPENDIX 3

Patient information in Study IV [In Danish]

«Adressekartotek»

'Min Epilepsi' på Sundhed.dk

Dato 16-06-2016
Liv Schougaard
Tlf.: 7843 3537
Mobil: 6179 2004
livschou@rm.dk

Kære «Navn»

Vi skriver til dig for at fortælle, at din kontakt til epilepsiambulatoriet nu overgår til et nyt og mere fleksibelt tilbud, hvor du selv styrer kontakten til ambulatoriet.

HVAD SKAL DU GØRE?
Læs mere i vedlagte information.

Via siden **Min Epilepsi** på Sundhed.dk, kan du nu fremover besvare et spørgeskema om din epilepsi, når du har behov for kontakt med epilepsiambulatoriet.

Din besvarelse sendes automatisk til ambulatoriet, og du bliver hurtigst muligt kontaktet af personalet i ambulatoriet.

Det er frivilligt, om du vil deltage i denne ordning eller om du gerne vil tilbage på ordningen, hvor vi sender spørgeskemaer til dig med faste mellemrum.

Hvis du **ikke** ønsker at styre kontakten til ambulatoriet selv eller hvis du har brug for mere information, bedes du kontakte Liv Schougaard.

Med venlig hilsen

Liv Schougaard
Sygeplejerske, ph.d.-studerende
Telefon: 78433537 eller 61792004
E-mail: livschou@rm.dk

Per Sidenius
Ledende overlæge
Neurologisk afdeling
Aarhus Universitetshospital

Har du spørgsmål om projektet?

Har du spørgsmål til projektet, er du meget velkommen til at kontakte:

Liv Schougaard
Projektansvarlig sygeplejerske, ph.d.-studerende

Telefon: 78433537 eller 61792004
E-mail: livschou@rm.dk

regionmidtjylland

Min Epilepsi

via sundhed.dk

Vil du gerne se dine spørgeskemabesvarelser?

Vil du selv styre din kontakt til epilepsiambulatoriet?

Deltagerinformation:

Du er udvalgt til at deltage i et projekt, hvor vi tilbyder et nyt og mere fleksibelt tilbud til dig.
Via siden **Min Epilepsi** kan du kontakte epilepsiambulatoriet ved at besvare et spørgeskema om din epilepsi. Du kan også se dine tidligere besvarelser af epilepsispørgeskemaet.
Du har adgang til internetsiden via Sundhed.dk

Læs mere på næste side

AmbuFlex

Hvad går projektet ud på?

Projektet går ud på at tilbyde dig et nyt og mere fleksibelt tilbud, hvor du selv i højere grad styrer kontakten til dit epilepsiambulatorium.

Du skal gå ind på ”Min Epilepsi” og besvare spørgeskemaet, hvis du ønsker at komme i kontakt med en sygeplejerske eller læge fra epilepsiambulatoriet. Det kan f.eks. være, hvis du oplever ændringer i dine symptomer.

Du bestemmer selv, hvornår du besvarer spørgeskemaet.

Din besvarelse sendes automatisk til epilepsiambulatoriet. En sygeplejerske eller læge fra ambulatoriet ser din besvarelse og kontakter dig enten per telefon eller tilbyder dig en tid i ambulatoriet.

Du er altid velkommen til at ringe til ambulatoriet, hvis du har brug for hurtig kontakt.

Hvis der går lang tid, uden at vi hører fra dig, får du automatisk et spørgeskema tilsendt.

Hvad sker der nu?

Du styrer nu selv kontakten til ambulatoriet, som beskrevet. Vejledning til internetsiden *”Min Epilepsi”* er vedlagt.

Projektet varer i 1 ½ år, men du kan til enhver tid vælge at træde ud af projektet.

Alle medarbejdere, der er involveret i projektet, har tavshedspligt, og oplysninger vil blive behandlet strengt fortroligt.

Projektet er godkendt af Datatilsynet.

Det er frivilligt, om du vil deltage eller ej.

Hvis du **ikke** ønsker at deltage, bedes du kontakte den projektansvarlige sygeplejerske, **Liv Schougaard:**

Telefon: 78433537 eller 61792004
E-mail: livschou@rm.dk

Gem denne folder!
Deltager du i projektet, er det vigtigt, at du gemmer denne folder, da den indeholder oplysninger, du kan få brug i projektets forløb.

Min Epilepsi

via sundhed.dk

Vejledning til at logge på ”Min Epilepsi” internetsiden via Sundhed.dk

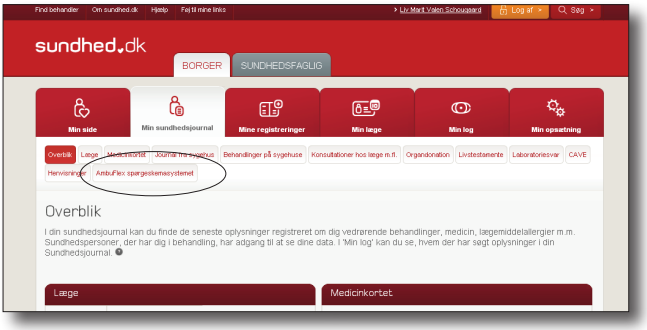
1. Gå ind på www.sundhed.dk

2. Log på med NEM ID

   Hvis det er første gang, du logger på sundhed.dk, skal du afgive et samtykke.

3. Gå ind under
   Min sundhedsjournal

4. Klik på
   AmbuFlex spørgeskemasystemet



5. Du kommer nu frem til siden “Min epilepsi”

   Tillad pop-up vinduer, hvis du ikke kan åbne siden

AmbuFlex

vend →

Min Epilepsi internetsiden består af fire dele:

1. Besvar spørgeskema: Her kan du besvare epilepsispørgeskemaet, når du har brug for at tale med en sygeplejerske eller læge.

   Du har også mulighed for at tilføje en kommentar til sidst i spørgeskemaet.

   Du vælger selv kontaktform.

2. Tidligere besvarelser: Her har oversigt over alle dine tidligere spørgeskemabesvarelser.

3. Information: Her kan du læse om internetsiden ”Min Epilepsi”, epilepsispørgeskemaet og hvorfor der spørges ind til forskellige emner.

4. Kontaktoplysninger på ambulatoriet: Her kan du finde mailadresse og telefonnummer til epilepsiambulatoriet.

AmbuFlex

APPENDIX 4

Co-authorship declarations

AARHUS UNIVERSITY

Declaration of co-authorship concerning article for PhD dissertations

Full name of the PhD student: Liv Marit Valen Schougaard

This declaration concerns the following article/manuscript:

Title:	Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study
Authors:	Schougaard LMV, de Thurah A, Christensen J, Lomborg K, Maindal HT, Mejdahl CT, Vestergaard JM, Winding TN, Biering K, Hjollund NH

The article/manuscript is: Published ☒ Accepted ☐ Submitted ☐ Ready for submission ☐

If published, state full reference: Schougaard LMV, de Thurah A, Christensen J, Lomborg K, Maindal HT, Mejdahl CT, Vestergaard JM, Winding TN, Biering K, Hjollund NH. Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study. Qual Life Res 2020 Jan 3

If accepted or submitted, state journal:

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No ☒ Yes ☐ If yes, give details:

Your contribution

Please rate (A-F) your contribution to the elements of this article/manuscript, **and** elaborate on your rating in the free text section below.

A. Has essentially done all the work (>90%)
B. Has done most of the work (67-90 %)
C. Has contributed considerably (34-66 %)
D. Has contributed (10-33 %)
E. No or little contribution (<10%)
F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	C
Free text description of PhD students contribution ***(mandatory)*** NHH, AdT, HTM, and LMVS designed the study.	
The acquisition, analysis, or interpretation of data:	B
Free text description of PhD students contribution ***(mandatory)*** NHH, AdT, HTM, JC, JMV, TNW, KB, and LMVS were involved in data collection, management, plan for statistical analyses, or interpretation of data. LMVS performed the statistical analyses.	
Drafting the manuscript:	A
Free text description of PhD students contribution ***(mandatory)*** LMVS drafted the manuscript. All authors contributed to critical revision of the manuscript and approved the final manuscript.	

AARHUS
UNIVERSITY

Submission process including revisions:	A
Free text description of PhD students contribution ***(mandatory)*** ***LMVS submitted the final manuscript. LMVS drafted a revised manuscript after peer review. All authors contributed to critical revision of the revised manuscript and approved the final revised manuscript. LMVS submitted the revision.***	

Signatures of first- and last author, and main supervisor

Date	Name	Signature
2020/01/13	Liv Marit Valen Schougaard	
2020/01/13	Niels Henrik Hjøllund	
2020/01/13	Niels Henrik Hjøllund	

Date: 13/1 - 2020

Signature of the PhD student

AARHUS UNIVERSITY

Declaration of co-authorship concerning article for PhD dissertations

Full name of the PhD student: Liv Marit Valen Schougaard

This declaration concerns the following article/manuscript:

Title:	Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test-retest reliability study
Authors:	Schougaard LMV, de Thurah A, Christiansen DH, Sidenius P, Hjollund NH

The article/manuscript is: Published ☒ Accepted ☐ Submitted ☐ Ready for submission ☐

If published, state full reference: Schougaard LMV, de Thurah A, Christiansen DH, Sidenius P, Hjollund NH. Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test-retest reliability study. BMJ Open 2018 Jul 25;8(7):e021337-2017-021337

If accepted or submitted, state journal:

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No ☒ Yes ☐ If yes, give details:

Your contribution

Please rate (A-F) your contribution to the elements of this article/manuscript, **and** elaborate on your rating in the free text section below.

A. Has essentially done all the work (>90%)
B. Has done most of the work (67-90 %)
C. Has contributed considerably (34-66 %)
D. Has contributed (10-33 %)
E. No or little contribution (<10%)
F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	C
Free text description of PhD students contribution ***(mandatory)*** NHH and LMVS designed the study in collaboration with PS, AdT, and DHC	
The acquisition, analysis, or interpretation of data:	B
Free text description of PhD students contribution ***(mandatory)*** LMVS participated in recruitment of participants, data collection, and performed the statistical analyses. NHH, AdT, DHC and LMVS contributed to interpretation of data.	
Drafting the manuscript:	A
Free text description of PhD students contribution ***(mandatory)*** LMVS drafted the manuscript. All authors contributed to critical revision of the manuscript and approved the final manuscript.	
Submission process including revisions:	A

AARHUS UNIVERSITY

Free text description of PhD students contribution ***(mandatory)***
LMVS submitted the final manuscript. LMVS drafted a revised manuscript after peer review. All authors contributed to critical revision of the revised manuscript and approved the final revised manuscript. LMVS submitted the revision.

Signatures of first- and last author, and main supervisor

Date	Name	Signature
2020/01/13	Liv Marit Valen Schougaard	
2020/01/13	Niels Henrik Hjøllund	
2020/01/13	Niels Henrik Hjøllund	

Date: 13/1 - 2020

Liv M. O. Schougaard

Signature of the PhD student

AARHUS UNIVERSITY

Declaration of co-authorship concerning article for PhD dissertations

Full name of the PhD student: Liv Marit Valen Schougaard

This declaration concerns the following article/manuscript:

Title:	Test-retest reliability and measurement error of the Danish WHO-5 Well-being Index in outpatients with epilepsy
Authors:	Schougaard LMV, de Thurah A, Bech P, Hjollund NH, Christiansen DH

The article/manuscript is: Published ☒ Accepted ☐ Submitted ☐ Ready for submission ☐

If published, state full reference: Test-retest reliability and measurement error of the Danish WHO-5 Well-being Index in outpatients with epilepsy. Health Qual Life Outcomes 2018 16:175-018-1001-0

If accepted or submitted, state journal:

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No ☒ Yes ☐ If yes, give details:

Your contribution

Please rate (A-F) your contribution to the elements of this article/manuscript, **and** elaborate on your rating in the free text section below.

A. Has essentially done all the work (>90%)
B. Has done most of the work (67-90 %)
C. Has contributed considerably (34-66 %)
D. Has contributed (10-33 %)
E. No or little contribution (<10%)
F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	C
Free text description of PhD students contribution ***(mandatory)*** The research team including DHC, NHH, AdT and LMVS conceived the study.	
The acquisition, analysis, or interpretation of data:	B
Free text description of PhD students contribution ***(mandatory)*** The research team conveived a statistical analysis plan and interpretation of data. LMVS participated in data collection and performed the statistical analyses.	
Drafting the manuscript:	A
Free text description of PhD students contribution ***(mandatory)*** LMVS drafted the manuscript. All authors contributed to critical revision of the manuscript and approved the final manuscript.	
Submission process including revisions:	A

AARHUS UNIVERSITY

Free text description of PhD students contribution ***(mandatory)***
LMVS submitted the final manuscript. LMVS drafted a revised manuscript after peer review. All authors contributed to critical revision of the revised manuscript and approved the final revised manuscript. LMVS submitted the revision.

Signatures of first- and last author, and main supervisor

Date	Name	Signature
2020/01/13	Liv Marit Valen Schougaard	
2020/01/10	David Høyrup Christiansen	
2020/01/13	Niels Henrik Hjøllund	

Date: 13/1-2020

Signature of the PhD student

AARHUS UNIVERSITY

Declaration of co-authorship concerning article for PhD dissertations

Full name of the PhD student: Liv Marit Valen Schougaard

This declaration concerns the following article/manuscript:

Title:	Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: study protocol for a pragmatic randomised controlled study
Authors:	Schougaard LMV, Mejdahl CT, Petersen KH, Jessen A, de Thurah A, Sidenius P, Lomborg K, Hjollund NH

The article/manuscript is: Published ☒ Accepted ☐ Submitted ☐ Ready for submission ☐

If published, state full reference: Schougaard LMV, Mejdahl CT, Petersen KH, Jessen A, de Thurah A, Sidenius P, Lomborg K, Hjollund NH. Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: study protocol for a pragmatic randomised controlled study. BMC Health Serv Res 2017 17:83-017-2015-8

If accepted or submitted, state journal:

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No ☒ Yes ☐ If yes, give details:

Your contribution

Please rate (A-F) your contribution to the elements of this article/manuscript, **and** elaborate on your rating in the free text section below.

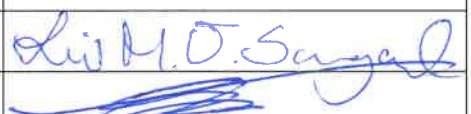
A. Has essentially done all the work (>90%)
B. Has done most of the work (67-90 %)
C. Has contributed considerably (34-66 %)
D. Has contributed (10-33 %)
E. No or little contribution (<10%)
F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	C
Free text description of PhD students contribution ***(mandatory)*** NHH and LMVS conceived the study in collaboration with PS, AdT, and KL.	
The acquisition, analysis, or interpretation of data:	F
Free text description of PhD students contribution ***(mandatory)*** The study is a protocol paper of a pragmatic randomised controlled study. No analysis or interpretation of data were performed.	
Drafting the manuscript:	A
Free text description of PhD students contribution ***(mandatory)*** LMVS drafted the manuscript. All authors contributed to critical revision of the manuscript and approved the final manuscript.	

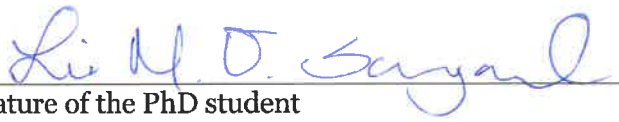
AARHUS UNIVERSITY

Submission process including revisions:	A
Free text description of PhD students contribution ***(mandatory)*** ***LMVS submitted the protocol paper and responded to the comments from the Editor. A less detailed protocol was reviewed by one of the Funding sources, thus, the protocol did not undergo peer review. The review comments from the Funding Source (TrygFonden) was sent to the Editor.***	

Signatures of first- and last author, and main supervisor

Date	Name	Signature
2020.01.13	Liv Marit Valen Schougaard	
2020.01.13	Niels Henrik Hjøllund	
2020.01.13	Niels Henrik Hjøllund	

Date: 13/1 - 2020



Signature of the PhD student

AARHUS UNIVERSITY

Declaration of co-authorship concerning article for PhD dissertations

Full name of the PhD student: Liv Marit Valen Schougaard

This declaration concerns the following article/manuscript:

Title:	Patient-initiated versus fixed-interval patient-reported outcome-based follow-up in outpatients with epilepsy: a pragmatic randomized controlled trial
Authors:	Schougaard LMV, Mejdahl CT, Christensen J, Lomborg K, Maindal HT, de Thurah A, Hjollund NH

The article/manuscript is: Published ☒ Accepted ☐ Submitted ☐ Ready for submission ☐

If published, state full reference: Schougaard LMV, Mejdahl CT, Christensen J, Lomborg K, Maindal HT, de Thurah A, Hjollund NH. Patient-initiated versus fixed-interval patient-reported outcome-based follow-up in outpatients with epilepsy: a pragmatic randomized controlled trial. J Patient Rep Outcomes 2019 Sep 13;3(1):61-019-0151-0

If accepted or submitted, state journal:

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No ☒ Yes ☐ If yes, give details:

Your contribution

Please rate (A-F) your contribution to the elements of this article/manuscript, **and** elaborate on your rating in the free text section below.

A. Has essentially done all the work (>90%)
B. Has done most of the work (67-90 %)
C. Has contributed considerably (34-66 %)
D. Has contributed (10-33 %)
E. No or little contribution (<10%)
F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	C
Free text description of PhD students contribution ***(mandatory)*** NHH and LMVS conceived the study in collaboration with PS, AdT, and KL.	
The acquisition, analysis, or interpretation of data:	B
Free text description of PhD students contribution ***(mandatory)*** A statistical analysis plan was conceived by the research team and discussed with a statistician. LMVS participated in data collection and performed the statistical analyses.	
Drafting the manuscript:	A
Free text description of PhD students contribution ***(mandatory)*** LMVS drafted the manuscript. All authors contributed to critical revision of the manuscript and approved the final manuscript.	

AARHUS
UNIVERSITY

Submission process including revisions:	A
Free text description of PhD students contribution ***(mandatory)*** ***LMVS submitted the final manuscript. LMVS drafted a revised manuscript after peer review. All authors contributed to critical revision of the revised manuscript and approved the final revised manuscript. LMVS submitted the revision.***	

Signatures of first- and last author, and main supervisor

Date	Name	Signature
2020/01/13	Liv Marit Valen Schougaard	
2020/01/13	Niels Henrik Hjøllund	
2020/01/13	Niels Henrik Hjøllund	

Date: 13/1-2020

Signature of the PhD student