



Statistical Analysis Plan

Q.RARE.LI

Improving health-related quality of life in patients with rare autoimmune liver diseases by structured peer-delivered support: a transnational effectiveness-implementation hybrid trial

Trial registration: <https://doi.org/10.1186/ISRCTN15030282>

Version: 1.0

Date: November 22, 2024

Author: Dr. Natalie Uhlenbusch

Reviewers: Prof. Dr. Bernd Löwe, Prof. Dr. Antonia Zapf

Institution: UKE



Document:	Statistical Analysis Plan for Final Analysis
Sponsor Project Code:	Q.RARE.LI
Study No.	Trial registration: https://doi.org/10.1186/ISRCTN15030282
Study short title:	Improving health-related quality of life in patients with rare autoimmune liver diseases by structured peer-delivered support: a transnational effectiveness-implementation hybrid trial
Version / date:	1.0 / November 22, 2024
Author:	Dr. Natalie Uhlenbusch
Sponsor:	UKE
Statistician:	Prof. Dr. Antonia Zapf Department of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
Study coordinators:	Prof. Dr. Bernd Löwe, Dr. Natalie Uhlenbusch, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
Project partners:	Prof. Dr. Gideon Hirschfield, Division of Gastroenterology and Hepatology, University Health Network, University of Toronto, Toronto, Canada Prof. Dr. Anja Geerts, Department of Internal Medicine and Pediatrics, Faculty of Medicine and Health Sciences, Ghent University Hospital, Gent, Belgium Prof. Dr. Piotr Milkiewicz, Liver and Internal Medicine Unit, Medical University of Warsaw, Warsaw, Poland Prof. Dr. Mária Papp, Department of Gastroenterology, Institute of Internal Medicine, University of Debrecen, Debrecen, Hungary

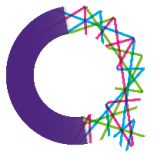
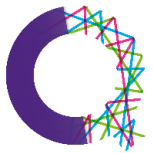
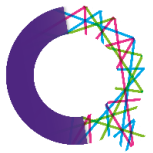


Table of Contents

Version history	5
Signature Page	6
Abbreviations	7
1 Introduction	8
1.1 Background and rationale	8
1.2 Objectives	9
2 Study Methods	9
2.1 Trial design	9
2.2 Randomization	11
2.3 Sample size	11
2.4 Framework	11
2.5 Statistical interim analyses and stopping guidance	11
2.6 Timing of final analysis	12
2.7 Timing of outcome assessments	12
3 Statistical Principles	12
3.1 Confidence intervals and <i>P</i> values	12
3.2 Adherence and protocol deviations	12
3.3 Analysis populations	12
3.3.1 Intention to treat (ITT) population	12
3.3.2 Per Protocol (PP) population	13
4 Trial Population	13
4.1 Screening	13
4.2 Eligibility	13
4.3 Recruitment	13
4.4 Withdrawal/follow-up	13
4.5 Baseline patient characteristics	14



5	Analysis	14
5.1	Outcome definitions	14
5.1.1	Effectiveness outcomes	14
5.1.2	Implementation outcomes	14
5.2	Analysis methods	15
5.2.1	Effectiveness outcomes	15
5.2.2	Implementation outcomes	15
5.3	Missing data.....	15
5.4	Additional analyses	15
5.5	Harms.....	16
5.6	Statistical software.....	16
6	References	17



Version history

Version	Date	Section number changed	Description/Changes made
0.1	20-NOVEMBER- 2024		First draft
1.0	22-NOVEMBER- 2024		Finalized



Signature Page

Prof. Dr. Bernd Löwe

Principal Investigator, sponsor delegated person

Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

16.12.2024

Date

Signature

Prof. Dr. Antonia Zapf

Trial Statistician

Department of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

6.12.2024

Date

Signature

Dr. Natalie Uhlenbusch

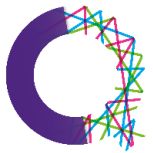
Principal investigator, SAP Author

Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

22.11.2024

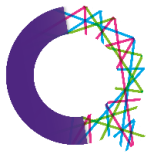
Date

Signature



Abbreviations

AIH	Autoimmune hepatitis
CAU	Care-as-usual
DSMB	Data Safety and Monitoring Board
ITT	Intention-to-treat
NRS	Numeric rating scale
PBC	Primary biliary cholangitis
PSC	Primary sclerosing cholangitis
PP	Per Protocol
RCT	Randomised Controlled Trial
SAP	Statistical Analysis Plan
SD	Standard Deviation



1 Introduction

This Statistical Analysis Plan (SAP) is based on the published study protocol (1) and follows the guideline for statistical analysis plans by Gamble and colleagues (2017 (2)). Some points of the statistical methods and of the study design are already described in the study protocol. This SAP aims to further specify the procedures and statistical methods applied during the final analysis of the study data.

1.1 Background and rationale

In absence of a cure for most rare conditions, improving patients' quality of life is a central healthcare aim. Psychosocial support alongside the treatment of somatic symptoms in order to help patients adjust to their disease and stay mentally healthy is key to this (3). Support programs for patients with common chronic diseases do not adequately address the unique challenges of patients with rare diseases, i.e. limited access to adequate diagnosis, treatment, and information about the diseases, a lack of contact to peers with the same condition (4-6). Moreover, the high number of various rare diseases makes isolated consideration of individual conditions in psychosocial care unfeasible (7). Patients' geographic dispersion presents an additional difficulty because it limits their access to healthcare options. Consequently, psychosocial support needs of patients with rare diseases are insufficiently met.

To address these needs, we developed a location-independent, trans-diagnostic and peer-delivered psychosocial support program specifically for individuals with rare diseases (8). The program combines manual-based self-help and structured telephone-based peer-support. It is based on Acceptance and Commitment Therapy and participants complete it from home over the course of six weeks. Peer-supporters receive training, counselling guidelines and supervision. A more detailed description of the intervention is published within the study protocol (1). We demonstrated the efficacy and acceptability of the intervention in a first randomized controlled trial (RCT) including patients with four different rare diseases (9). However, the effectiveness of this intervention in routine care has not yet been demonstrated and it has not been implemented in any national health system. This is the starting point of Q.RARE.LI. We will assess the effectiveness of our peer-delivered psychosocial support intervention in routine care for patients with rare autoimmune liver diseases and prepare the implementation on an international level at five different sites (Hamburg, Germany; Toronto, Canada; Gent, Belgium; Warsaw, Poland; Debrecen, Hungary). Our long-term aim is to establish the intervention internationally, making it widely accessible to patients with rare diseases in order to improve quality of life.



1.2 Objectives

The primary research objective of Q.RARE.LI is to investigate the effectiveness of the peer-delivered psychosocial support intervention under routine care conditions in five different healthcare settings and to prepare its implementation. We therefore have two research foci:

A. Effectiveness: Objective: To investigate the effectiveness of the intervention in routine health care in terms of health-related quality of life, depression and anxiety severity, illness acceptance, perceived helplessness, social support, and self-management abilities. Primary hypothesis: Under routine care conditions, structured peer-delivered psychosocial support in addition to CAU leads to better mental health-related quality of life compared to CAU alone in patients with rare liver diseases at post-assessment. Secondary hypotheses: Under routine care conditions, structured peer-delivered psychosocial support in addition to CAU leads to better mental health-related quality of life at 3-month follow-up and to improved outcomes regarding a) physical health-related quality of life, b) depression severity, c) anxiety severity, d) illness acceptance, e) perceived helplessness f) social support and g) self-management abilities compared to CAU alone in patients with rare liver diseases both at post-assessment and 3-month follow-up.

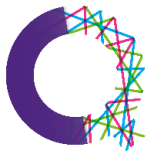
B. Implementation: Objective: To assess implementation outcomes, examine and prepare the implementation context, and develop implementation strategies in each of the five participating countries. Primary hypothesis: The intervention shows high acceptability and feasibility in routine care as indicated by: 1) >75% of the patients completing the intervention, 2) >75% of the patients rating the intervention as a) helpful and b) appropriate, 3) >75% of the stakeholders delivering the intervention rating it as a) feasible, b) appropriate and c) helpful for patients. We will also identify implementation barriers and facilitators.

2 Study Methods

2.1 Trial design

We conduct an effectiveness-implementation hybrid trial, a study design blending the processes of effectiveness and implementation research in order to better understand the contextual factors related to the success of interventions (10, 11). An overview is displayed in **Figure 1**.

Effectiveness of the intervention is evaluated within a two-armed RCT comparing structured and peer-delivered psychosocial support plus CAU to CAU alone in adult patients with autoimmune rare liver diseases. We include $N=240$ patients with autoimmune hepatitis (AIH), primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC) who are randomly assigned to either an intervention or a



control group. The intervention group participates in the program after the baseline assessment in addition to their care-as-usual (CAU). The control group receives CAU alone, but has the opportunity to participate in the program after their last assessment. The intervention is applied in each of the five participating sites under country-specific routine care conditions. Outcomes are assessed at three time points (baseline, post, 3-month follow-up). The primary effectiveness outcome is mental health-related quality of life (SF-12 (12)).

Implementation outcomes are assessed within a mixed-methods process evaluation. We conduct a quantitative, cross-sectional survey including patients, peer-counsellors and healthcare providers that completed the intervention or were involved in its delivery. The main quantitative implementation outcomes are perceived acceptability and feasibility of the intervention. In addition, we conduct a qualitative study after the last quantitative assessment is completed, including four different stakeholder groups: patients, peer-counsellors, healthcare providers, healthcare leaders (e.g. health insurers). Primary aim of the qualitative study is to better understand the implementation context in each country and identify barriers and facilitators for a successful and sustainable implementation. Based on the results, each country derives appropriate implementation strategies and develop a concrete implementation plan.

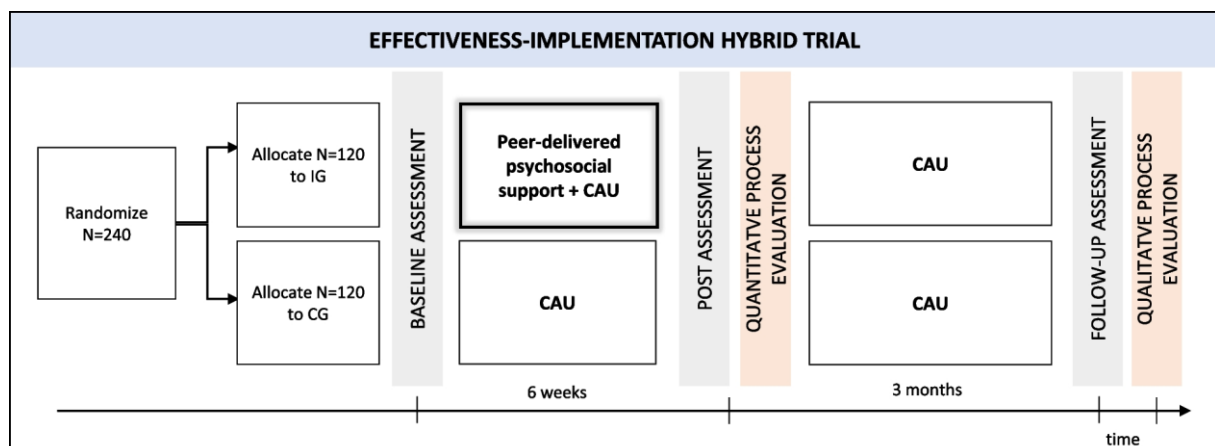


Figure 1. Study design.

Each partner obtained a positive ethics vote by an independent national ethics committee. The trial is conducted in accordance with the Declaration of Helsinki, guidelines for Good Clinical Practice, national and local laws. Before inclusion, eligible participants are informed about the course of the study verbally and in written form and provide written informed consent. The data is stored in pseudonymized form. Any changes to the study protocol are listed in the study registry and publications.



2.2 Randomization

A fixed randomization schedule with an allocation ratio 1:1, stratified by country, was deposited in the REDCap system. The randomization table was created and deposited by a researcher who is not involved in the project. Patients are electronically randomized according to this schedule after signing the informed consent form. They are informed about their group allocation after the baseline assessment. The researchers involved in Q.RARE.LI do not have any influence on the randomization process.

2.3 Sample size

Our primary effectiveness outcome is the group difference in the baseline-adjusted mean score in mental health-related quality of life at post-assessment, assessed with the mental component of the Short-Form Health Survey (SF-12 (12)). Taking the results of our prior RCT into account (9), we assume a between groups effect size of $d=0.4$. Based on two-sided testing with $\alpha=0.05$ and $1-\beta=0.8$, $N=100$ patients are needed per group (intention-to-treat analysis), yielding in $N=200$ to be analyzed. Considering our prior RCT (9), we conservatively expect an inclusion rate of 60% and approximately 15% loss to follow-up. Thus, $N=240$ patients are included at the five study sites over 15 months.

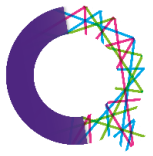
For the quantitative implementation part of the trial, all patients in the intervention group ($N=100$), all peer-counselors ($\sim N=40$) and all involved healthcare providers ($\sim N=25$) are invited to participate in the survey. With an estimated response rate of 80% (9), we expect to include $N=132$ participants across all sites. For the qualitative focus groups, at least $N=5$ participants per stakeholder group (patients, peer-counsellors, healthcare providers, healthcare leaders) and site are invited to take part, resulting in an estimated sample size of $N=100$ in total.

2.4 Framework

Q.RARE.LI is a randomized superiority trial with two study arms, aiming to investigate whether structured and peer-delivered psychosocial support in addition to CAU leads to improved mental quality of life compared to CAU alone in patients with rare autoimmune liver diseases.

2.5 Statistical interim analyses and stopping guidance

No interim analyses will be conducted. Adverse events are monitored and reported to the Data Safety and Monitoring Board (DSMB). Major events which need to be monitored comprise acute suicidality, suicidal acts, and life-threatening deterioration of health status. For the individual patient, the trial procedure stops if any adverse events or withdrawal of informed consent occur. The whole trial is discontinued if the project consortium or the DSMB detect significant associations between study participation and adverse events or a differential association between the experimental condition and adverse events.



2.6 Timing of final analysis

The data analysis starts after the last follow-up assessment is completed. The final analysis of the primary outcome takes place after the database has been reviewed for completeness and accuracy and after data cleaning is completed (estimated time point: December 2024).

2.7 Timing of outcome assessments

The primary and secondary outcomes within the RCT are assessed at baseline, after approximately 7 weeks (post-intervention) and at a 3-month follow-up. The quantitative data of the mixed-methods process evaluation are assessed at the post-assessment. The qualitative data of the mixed-methods study are collected after the follow-up assessment.

3 Statistical Principles

3.1 Confidence intervals and *P* values

All applicable statistical tests will be two-sided and are performed using a 5% significance level. Analyses of secondary outcomes will be performed exploratory, without adjustment for multiplicity. All confidence intervals presented will be 95% and two-sided.

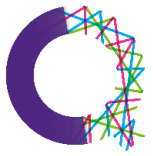
3.2 Adherence and protocol deviations

Peer-counsellors write a brief summary of each counselling session indicating whether the patient worked through the respective chapter of the self-help book (binary item: yes/no) and how much the session was in line with the counselling guidelines (numeric rating scale (NRS) from 0=not at all to 10=fully in line with the guidelines). In addition, peer-counsellors inform us if a session was skipped or if the patient wants to abort the program. In case a session was skipped or the patients did not work through the chapter, participants have the opportunity to make up the session at a later time. The total duration of the program participation is documented by peer-counsellors as well. A protocol deviation occurs if patients missed 15% or more of the program, which means that they have to complete at least 5 out of 6 sessions (chapters and counselling).

3.3 Analysis populations

3.3.1 Intention to treat (ITT) population

The primary analysis population is the intention-to-treat (ITT) population. All patients for whom data is available from at least one follow-up measurement point (full analysis set) will be included in assessment of the primary outcome in accordance with the intention-to-treat principle.



3.3.2 Per Protocol (PP) population

The PP population, including all patients randomized who do not have a protocol violation (definition in section 3.2) and for whom data is available from at least one follow-up measurement point (full analysis set), will be included in the per protocol analyses. The PP population will be additionally analyzed as part of a sensitivity analysis.

4 Trial Population

4.1 Screening

Patients' eligibility is assessed in structured telephone-based screening interviews. The number of conducted interviews and the reason for exclusion are assessed and will be reported as part of the CONSORT diagram (see also 4.3)

4.2 Eligibility

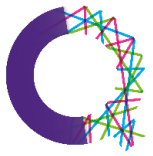
Inclusion criteria are the diagnosis of a rare autoimmune rare liver disease (AIH, PBC, PSC), a subjective psychosocial support need, an age of at least 18 years, and written informed consent. Exclusion criteria are a life-threatening health-status, acute suicidality, ongoing psychotherapy, severe cognitive, auditory or visual impairment and inability to complete data assessments.

4.3 Recruitment

The following information will be reported as part of the CONSORT diagram: number of patients screened, number of patients excluded (with reasons), number of patients included, number of patients randomized to each group, number of patients completing the intervention, number of patients completing post- and follow-up assessments, number of patients analyzed. The number of patients included per country will also be reported.

4.4 Withdrawal/follow-up

Patients are officially included in the study after they signed the informed consent form. Patients withdrawing their informed consent before completing the baseline assessment count as withdrawal. The number will be reported. Patients who withdraw from the study after the baseline assessment count as drop-outs. The number will also be reported and (if possible) the reason will be named. The lost-to-follow up rate is the proportion of patients who completed the baseline assessment but not the primary outcome assessment. We expect a lost-to-follow-up rate of 15%. In addition, we will report the number of patients in the intervention group who started the program but did not finish it and the time point at which they dropped out. Where available, we will state the reason for the earlier program termination.



4.5 Baseline patient characteristics

Baseline data includes demographical and clinical characteristics as well as primary and secondary outcomes. Categorical data will be summarized by numbers and percentages. Continuous data will be reported by mean, standard deviation (SD), and range (minimum, maximum). The number and percentage of missing observations will be presented for each variable. We will further assess treatment expectations (Treatment Expectation Questionnaire, TEX-Q (13)).

5 Analysis

5.1 Outcome definitions

5.1.1 Effectiveness outcomes

Primary outcome

The primary effectiveness outcome for this study is the baseline adjusted group difference in mental health-related quality of life at post-assessment. Mental health-related quality of life is measured with the mental component of the Short Form Health Survey (SF-12 (12)). The SF-12 measures psychological and physical aspects of generic, health-related quality of life. It is based on the 36-item version SF-36. Multiple studies demonstrated its sound psychometric properties (14, 15) and it has been used in a variety of chronic conditions, including patients with autoimmune liver diseases.

Secondary outcomes

Secondary outcomes are the baseline-adjusted group differences in mental health-related quality of life at follow-up as well as in physical health-related quality of life (SF-12), somatic symptom severity (Patient Health Questionnaire-15, PHQ-15 (16), depression severity (Patient Health Questionnaire-9, PHQ-9 (17, 18), anxiety severity (Generalized Anxiety Disorder Scale-7, GAD-7 (19)), illness cognitions (disease acceptance, helplessness, perceived benefits; Illness Cognition Questionnaire, ICQ (20)), social support (Social Support Questionnaire, F-SOZU (21)), and self-management abilities (Appraisal of Self-Care Agency Scale Revised, ASAS-R (22) at post- and follow-up assessment. We will also measure psychological burden related to somatic symptoms or associated health concerns (Somatic Symptom Disorder – B Criteria Scale, SSD-12 (23), general self-efficacy (Self-efficacy scale, SWE (24)), and illness perceptions (Brief Illness Perception Questionnaire, B-IPQ (25)).

5.1.2 Implementation outcomes

Quantitative implementation outcomes are assessed at post-assessment. We ask patients and peer-counsellors to evaluate the program and measure acceptability and feasibility on numeric rating scales from 0–10. We further ask healthcare providers to evaluate the program. Qualitative data are assessed in focus groups after the follow-up assessment is completed. We will assess perceived implementability



of the program into routine care with a standardized semi-structured interview guide. The development of this guide will be supported by the Consolidated Framework for Implementation Research (CFIR) Interview Guide Tool (CFIR Booklet (cfirguide.org)).

5.2 Analysis methods

5.2.1 Effectiveness outcomes

Primary outcome

To test our primary effectiveness hypothesis, we will perform a mixed linear model (covariance type AR1) using the restricted maximum likelihood method to produce estimates. The model will include time, country, and group as fixed factors, time as repeated effect, and the time×group interaction. We will include the baseline outcome score as a covariate and include a random intercept to model interindividual differences. Model specifications will not be adjusted based on model fit parameters. To assess group differences for each time point (post-assessment primary), we will determine the estimated marginal mean values. Effect sizes will be calculated by dividing the adjusted group mean difference by the observed SD of the total sample at baseline. We will check model assumptions by plotting residuals against estimated values and residual distribution against normal distribution. In case model assumptions are violated, we will produce robust estimates using Bootstrapping. As a sensitivity analysis, we will perform the data analysis with and without Bootstrapping.

Secondary outcomes

Secondary outcomes will be analyzed in the same manner as the primary outcome. Secondary outcomes are not corrected for multiple testing and should be considered exploratory.

5.2.2 Implementation outcomes

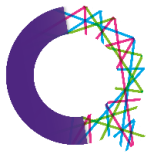
Quantitative data will be analyzed descriptively (mean, SD, range, percentages, figures). Qualitative data will be analyzed with qualitative content analysis. We will inductively identify topics in the data and deductively derive implementation barriers and facilitators based on the CFIR (www.cfirguide.org), i.e. considering the outer and inner setting, intervention characteristics, individuals involved, and the process. Based on these contextual conditions, we will derive implementation strategies with the CFIR.

5.3 Missing data

Outcome data will be multiply imputed (using MICE) if more than 5% are missing. In accordance with White et al. (2011 (26)), the number of imputations will be chosen depending on the proportion of missing data. As a sensitivity analysis, we will perform the data analysis with and without imputation.

5.4 Additional analyses

While the main analysis is conducted with the ITT population, the same analyses will be repeated with the PP population as a sensitivity analysis. Further exploratory subgroup analyses are performed by



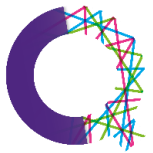
including the subgroup x group interaction into the model. A p-value of $<.15$ indicates a meaningful interaction and will lead to stratified analyses. Subgroup analyses will be performed for the following variable: gender (male, female), age (continuous for interaction), country (Belgium, Canada, Germany, Hungary Poland), diagnosis (PBS, PSC, AIH), degree of mental health-related quality of life impairment ($<\text{median}$, $\geq\text{median}$), treatment expectations (continuous for interaction), and whether the patient had a peer-counsellor with the same or a different rare disease (yes, no). We will further exploratively investigate outcomes in peer-counsellors.

5.5 Harms

In general, psychosocial support interventions bear no major risk for harm. In our first efficacy trial (9), no major adverse events occurred and no individual stated that the intervention harmed them. Rather, participating in the intervention was beneficial for patients. During the screening interviews, we explicitly ask whether patients have suicidal ideation. If this is detected, a site-specific algorithm is applied (e.g. contact the physician or consider psychiatric treatment). In case of acute suicidality, the patient is transferred to receive psychiatric treatment and excluded from the study. In addition, peer-counsellors are advised to report any harm they suspect in patients to their supervisor or the study coordinator at each site. If any harm is suspected, this is reported to the DSMB and an individual decision is made on whether participation should be stopped. Patients in the intervention group will be asked about any potential harm during the program evaluation at post-assessment. Suicidal ideation, attempts, suicide and any other potential harm are documented and reported.

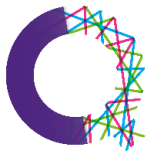
5.6 Statistical software

We will use SPSS 29.0 for all analyses.



6 References

1. Uhlenbusch N, Bal A, Balogh B, Braun A, Geerts A, Hirschfield G, et al. Improving quality of life in patients with rare autoimmune liver diseases by structured peer-delivered support (Q.RARE.LI): study protocol for a transnational effectiveness-implementation hybrid trial. *BMC Psychiatry*. 2023;23(1):193.
2. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017;318(23):2337-43.
3. Stanton AL, Revenson TA, Tennen H. Health psychology: psychological adjustment to chronic disease. *Annu Rev Psychol*. 2007;58:565-92. PMID: 16930096.
4. von der Lippe C, Diesen PS, Feragen KB. Living with a rare disorder: a systematic review of the qualitative literature. *Mol Genet Genomic Med*. 2017;5(6):758-73. PMID: 29178638.
5. Uhlenbusch N, Löwe B, Depping MK. Perceived burden in dealing with different rare diseases: a qualitative focus group study. *BMJ Open*. 2019;9(12):e033353. PMID: 31888936.
6. Jones DEJ, Sturm E, Lohse AW. Access to care in rare liver diseases: New challenges and new opportunities. *J Hepatol*. 2018;68(3):577-85. PMID: 29113911.
7. Eidt D, Frank M, Reimann A, Wagner TOF, Mittendorf T, Graf von der Schulenburg JM. Maßnahmen zur Verbesserung der gesundheitlichen Situation von Menschen mit seltenen Erkrankungen in Deutschland [Measures to improve the health situation of patients with rare diseases in Germany]. German Federal Ministry of Health; 2009.
8. Depping MK, Uhlenbusch N, Löwe B. Acceptance and commitment therapy meets peer-support. Development of a supportive self-care intervention for patients with rare diseases: A multi-stage development process. . *BMJ Open*. accepted.
9. Depping MK, Uhlenbusch N, Härter M, Schramm C, Löwe B. Efficacy of a brief, peer-delivered self-management intervention for patients with rare chronic diseases. *A Randomized Controlled Trial*. *JAMA Psychiatry*. 2021;78(6):607-15. PMID: 33625502.
10. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care*. 2012;50(3):217-26. PMID: 22310560.
11. Landes SJ, McBain SA, Curran GM. An introduction to effectiveness-implementation hybrid designs. *Psychiatry Res*. 2019;280:112513. PMID: 31434011.
12. Ware Jr JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-33. PMID: 8628042.
13. Shedden-Mora MC, Alberts J, Petrie KJ, Laferton JAC, von Blanckenburg P, Kohlmann S, et al. The Treatment Expectation Questionnaire (TEX-Q): Validation of a generic multidimensional scale measuring patients' treatment expectations. *PLoS One*. 2023;18(1):e0280472.
14. Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *Journal of Public Health*. 1997;19(2):179-86.
15. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51(11):1171-8.
16. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med*. 2002;64(2):258-66.
17. Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord*. 2004;81(1):61-6. PMID: 15183601.



18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13.
19. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092-7. PMID: 16717171.
20. Evers AWM, Kraaimaat FW, van Lankveld W, Jongen PJH, Jacobs JWG, Bijlsma JWJ. Beyond unfavorable thinking: The Illness Cognition Questionnaire for chronic diseases. *J Consult Clin Psychol*. 2001;69(6):1026-36. PMID: 11777106.
21. Fydrich T, Geyer M, Hessel A, Sommer G, Brähler E. Social support questionnaire (F-SozU): Norms of a representative sample. *Diagnostica*. 1999;45(4):212-16. WoS ID: 000083820700006.
22. Sousa VD, Zauszniewski JA, Bergquist-Beringer S, Musil CM, Neese JB, Jaber AF. Reliability, validity and factor structure of the Appraisal of Self-Care Agency Scale-Revised (ASAS-R). *J Eval Clin Pract*. 2010;16(6):1031-40.
23. Toussaint A, Löwe B, Brähler E, Jordan P. The Somatic Symptom Disorder - B Criteria Scale (SSD-12): Factorial structure, validity and population-based norms. *J Psychosom Res*. 2017;97:9-17.
24. Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. *Measures in health psychology: A user's portfolio Causal and control beliefs*. Windsor, UK: NFER-NELSON; 1995.
25. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res*. 2006;60(6):631-7.
26. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377-99.