LMU KLINIKUM DER UNIVERSITÄT MÜNCHEN
CAMPUS INNENSTADT
INSTITUT FÜR ALLGEMEINMEDIZIN

Statistical Analysis Plan

PICTURE – PTSD after ICU Survival

Caring for Patients with Traumatic Stress Sequelae following Intensive Medical Care

A multi-center, observer-blinded, randomized, controlled trial
with a psychological intervention

Principal Investigator („Leiter der klinischen Prüfung“):

Prof. Dr. Jochen Gensichen,

Institut für Allgemeinmedizin, Klinikum der Ludwig-Maximilians-Universität München

Nußbaumstr. 5, 80336 München

Responsible statistician:

Prof. Dr. Ulf-Dietrich Reips

Fachbereich Psychologie, Universität Konstanz

Universitätsstr. 10, 78646 Konstanz

DRKS-ID: DRKS00012589

ClinTrials.gov: NCT03315390

Version V 1.4, July 14th, 2023

DFG Grant: GE 2073/8-1

SIGNATURE PAGE

Principal Investigator _____

Jochen Gensichen (yyyy-mm-dd)

Statistician _____

Ulf-Dietrich Reips (yyyy-mm-dd)

Table of Contents

Table of Contents 3
1 Clinical Trial Synopsis 5
2 Schedule of Activities and Assessments 9
3 Trial Objectives and Endpoints 11
3.1 Add-on projects 11
3.2 Primary Objective 11
3.3 Primary Endpoint 11
3.4 Secondary Objectives 11
3.5 Secondary Endpoints 12
3.6 Additional secondary endpoints for “PICTURE-Economics” 12
3.7 Safety Variables 13
4 Trial Design 13
4.1 Number of subjects 13
4.2 Time Schedule 13
5 Trial Population, Eligibility Criteria, Recruitment 14
6 Randomization and Blinding 14
7 Trial Procedures 14
7.1 Methods of Assessment 14
7.1.1 PDS-5 total severity score (PTDS symptoms) 14
7.1.2 PHQ-9 (Depression) 14
7.1.3 OASIS (Anxiety) 15
7.1.4 Disability (WHODAS 2.0 – short version) 15
7.1.5 PAM-Score (Thirteen-Item Patient Activation Measure) 15
7.1.6 EQ-5D-5L (Quality of Life) 16
7.1.7 Questionnaires for the health economic evaluation 16
7.1.8 Covid-19 Burden Scale 16
7.2 Time schedule of Measurements 17
8 Statistical Methods 17
8.1 Planned Statistical Analyses 17
8.1.1 Analytical steps 17
8.1.2 Analysis of the primary efficacy endpoint 17
8.1.3 Analyses of secondary efficacy endpoints 18
8.1.4 Safety analysis 19

8.2 Interim analysis 20
8.3 Sample size calculation 20
8.4 Definition of populations included in the analyses 20
8.5 Protocol Violations 20
8.6 Handling of Drop-outs, Withdrawal, and Missing Data 20
9 Data Collection, Handling and Record Keeping 21
9.1 Data Forms and Data Entry 21
9.2 Data Transmission and Editing 21
10 Reporting 21
10.1 Statistical Report 21
11 Publication policy 21

1 Clinical Trial Synopsis

Clinical trial title	**PICTURE – PTSD aftEr ICU SURvival** Caring for Patients with Traumatic Stress Sequelae following Intensive Medical Care: A multi-center, observer-blinded, randomized, controlled trial with a psychological intervention
Trial short title	PICTURE
Trial Registry – nat. **international**	**DRKS-ID** (German Clinical Trials Register): DRKS00012589 **ClinicalTrials.gov:** NCT03315390
Medical Conditions	Post-traumatic stress disorder (PTSD) (ICD-10 F43.1; DSM-5 1.2.7) Post Intensive Care Syndrome (PICS) (Needham et al., 2012)
Interventions	**Experimental intervention:** A primary care version of a "Narrative Exposure Therapy" (**NET**-oriented, 3 sessions) delivered by the general practitioner (GP): *Session 1 (S1):* Diagnosis, psycho-education, and "lifeline", in which the patients constructs a chronology of their most significant life events *Session 2 (S2):* Narrative exposition, in which the patient recounts details of distressing situations that occurred in the Intensive Care Unit (ICU) *Session 3 (S3):* Narrative exposition of a stressful event, extracted from the patient's lifeline ("PDS event") *Telephone calls (TC) 1 – 7:* 7 telephone calls initiated by the GP practice affiliated medical assistant (MA) (S2, S3), including checking upon patients' well-being (PTSD-symptoms) and feedback to the GP. **Control intervention:** Improved treatment-as-usual (**iTAU**). GPs in the control group were instructed in evidence-based diagnosis and treatment of PTSD according to the S3-guideline. GPs contacted their patients for general checks and medical advice within 3 consultations (visit at GP's office).
Trial Population	Adult male and female post-ICU patients aged 18 to 85 years (both inclusive), showing symptoms of PTSD.

Trial Design	Prospective, national, multi-center, two-arm parallel-group (NET vs. iTAU), assessor-blinded, randomized controlled superiority trial with a psychological complex intervention delivered in the primary care setting
Trial Objectives	Primary objective: To demonstrate that the experimental intervention (NET) delivered by the GP is effective in reducing post-traumatic stress symptoms after intensive care measured by the PDS-5 total severity score, as compared to improved treatment as usual (iTAU) Secondary objectives: To demonstrate that NET is effective compared to iTAU in improving a) symptoms of depression and anxiety b) health-related quality of life, disability, and patient activation measure To demonstrate a favorable cost-effectiveness of NET compared to iTAU ("PICTURE-Economics") To describe the experiences of GP's in learning and implementing a psychotherapeutic treatment method in practice and how patients perceive the offer and the performance of a psychotherapeutic treatment delivered by their family physician ("PICTURE-Psychotherapy" add-on project) To check for the occurrence of adverse effects of the intervention
Trial Endpoints	**Primary Efficacy Endpoint:** • Posttraumatic Stress (Posttraumatic Diagnostic Scale for DSM-5): absolute change in PDS-5 total severity score from baseline to T1 (6 months after baseline) **Secondary Endpoints:** • Posttraumatic Stress: absolute change in PDS-5 total severity score from baseline to T2 (12 month after baseline) • Depression (PHQ-9 total score): absolute change from baseline to T1 and T2 • Anxiety (OASIS total score): absolute change from baseline to T1 and T2 • EQ-5D-5L: VAS at T1 and T2 • Disability (WHODAS 2.0 total score): absolute change from baseline to T1 and T2

	• Patient Activation Measure (PAM-13 total score): absolute change from baseline to T1 and T2 • "PICTURE-Economics": Cost-effectiveness at T1 and T2 – based on direct/ indirect costs as measured via modified CSSRI (Client Sociodemographic and Service Receipt Inventory) applied and QALYs (EQ-5D-5L index values) • "PICTURE-Psychotherapy" Qualitative aspects of attitudes and care of PTSD-patients in primary care
Subject Numbers	To be assessed for eligibility: **N = 3000 patients (GPs) in total** To be allocated to the trial: **N = 340 patients (GPs) in total will be randomized** (i.e. 170 per treatment arm) To be analyzed: **N = 318 patients (GPs) in total** (Recruitment concluded Dec 31st, 2022)
Trial Specific Measurements	Patient questionnaires (including mode of administration) • PDS-5 (Posttraumatic Stress): self-complete version on paper (by default at T0, T1, T2); or (telephone-) interview version (for non-responders at T0, T1, T2) • PHQ-9 (Depression) • OASIS (Anxiety) • EQ-5D-5L visual analogue scale (VAS); descriptive system: validated self-complete version on paper at T0, validated telephone interview version at T1, T2 • Disability (WHODAS 2.0, 12-item version) • PAM-13 (Patient Activation Measure) • modified version of CSSRI: interview version (assessed by the GP at T0; telephone interview version at T1 and T2)
Statistical Rationale	Primary efficacy analysis: The primary efficacy endpoint is the absolute change in PDS total severity score (Δ_6PDS) from baseline (T0) at month 6 (T1), no matter if derived from the PDS-5 questionnaire administered by mail or phone at T1. The combined null hypothesis for the primary efficacy endpoint at T1 is that distributions of absolute change score values are the same for NET and iTAU in patients with and without missing scores. Under the

	alternative hypothesis, we expect a shift in distributions with a clinically relevant standardized effect size in the order of 0.36 (Cohen's d). To test the null hypothesis of the confirmatory principal analysis, we will use a two-sample t-test assuming a normal distribution of the Δ_6PDS variable. The principal analysis will be performed according to the intention-to-treat (ITT) principle, unadjusted for baseline covariates or site. The significance level is set to alpha = 5% (two-sided). The confirmatory principal analysis will be verified using multivariate models adjusting for sociodemographic factors, drop-out and potential confounders. <u>Secondary endpoints and secondary analyses:</u> All secondary analyses will be exploratory, i.e. without adjustment for multiplicity, using adequate descriptive statistics as well as bivariate and multivariate statistical methods. The corresponding 95% confidence intervals for treatment group effects will be reported. <u>Safety analysis:</u> Safety analyses will be performed in the safety population. All observed safety events will be summarized using standard descriptive statistics stratified by the NET vs. iTAU conditions. <u>Health economic evaluation ("PICTURE-Economics"):</u> On the basis of the EQ-5D-5L index values and data reported by means of the modified CSSRI questionnaire, cost-effectiveness will be described by the incremental cost-effectiveness ratio (ICER), i.e. the ratio between the cost and effect differences between intervention and control group. To assess the uncertainty associated with the ICER, a series of net-benefit regressions will be performed, and a cost-effectiveness acceptability curve will be constructed. <u>Process evaluation ("PICTURE-Psychotherapy")</u> To investigate the subjective perception of the narrative intervention for patients and GPs respectively, as well as the perception of GP's in terms of delivering and usability of the NET-oriented therapy and the perception of patients in terms of relevance and impact, semi-structured interviews were conducted in a small subset of NET-participants with patients and GPs for qualitative analysis.

2 Schedule of Activities and Assessments

	STUDY PERIOD						
	Screening period		**Enrolment and allocation**		**Post-allocation**		
TIMELINE	**T -1** Screening	**2 weeks prior to T0** Re-Screening	**T 0**	**Randomi-zation** (≤ 2 weeks after T0)	***Intervention period*** *(3 visits)*	**T 1 (6 month after T0)**	**T2 (12 month after T0)**
Patient Informed Consent	X*		X				
Eligibility Screen	X	X					
Patient level: In-/Exclusion Criteria			X				
PC-PTSD-5	X	X					
SIS	X						
GP level: In-/Exclusion Criteria; GP characteristics		X					
GP Informed Consent		X					
ICU Data (incl. SOFA score)	X						
Patient medical history, demographics/ characteristics	X		X				
Covid-19 Burden Scale (started 03/20)			x			x	x
*Experimental intervention (**NET**)*				← - -	- - →		
*Control intervention (**iTAU**)*				← - -	- - →		
Documentation of SAE**				← - -	- - - -	- - - -	- - - →
PDS-5			X			X	X
PHQ-9			X			X	X
OASIS			X			X	X
EQ-5D-5L			X			X	X
PAM			X			X	X
WHODAS 2.0			X			X	X
Modified CSSRI (incl. concomitant (drug) therapy, sociodemographic data)			X			X	X
Documentation of concomitant medication / therapy			X				
Check of compliance, protocol adherence (patient and GP level)			X		X	X	X

*Patient Informed Consent for Screening; ** Safety documentation: to be performed by the treating GP starting at S1 (NET-condition).A deviation from the schedule of +/-3 weeks is acceptable for T1 and T2.

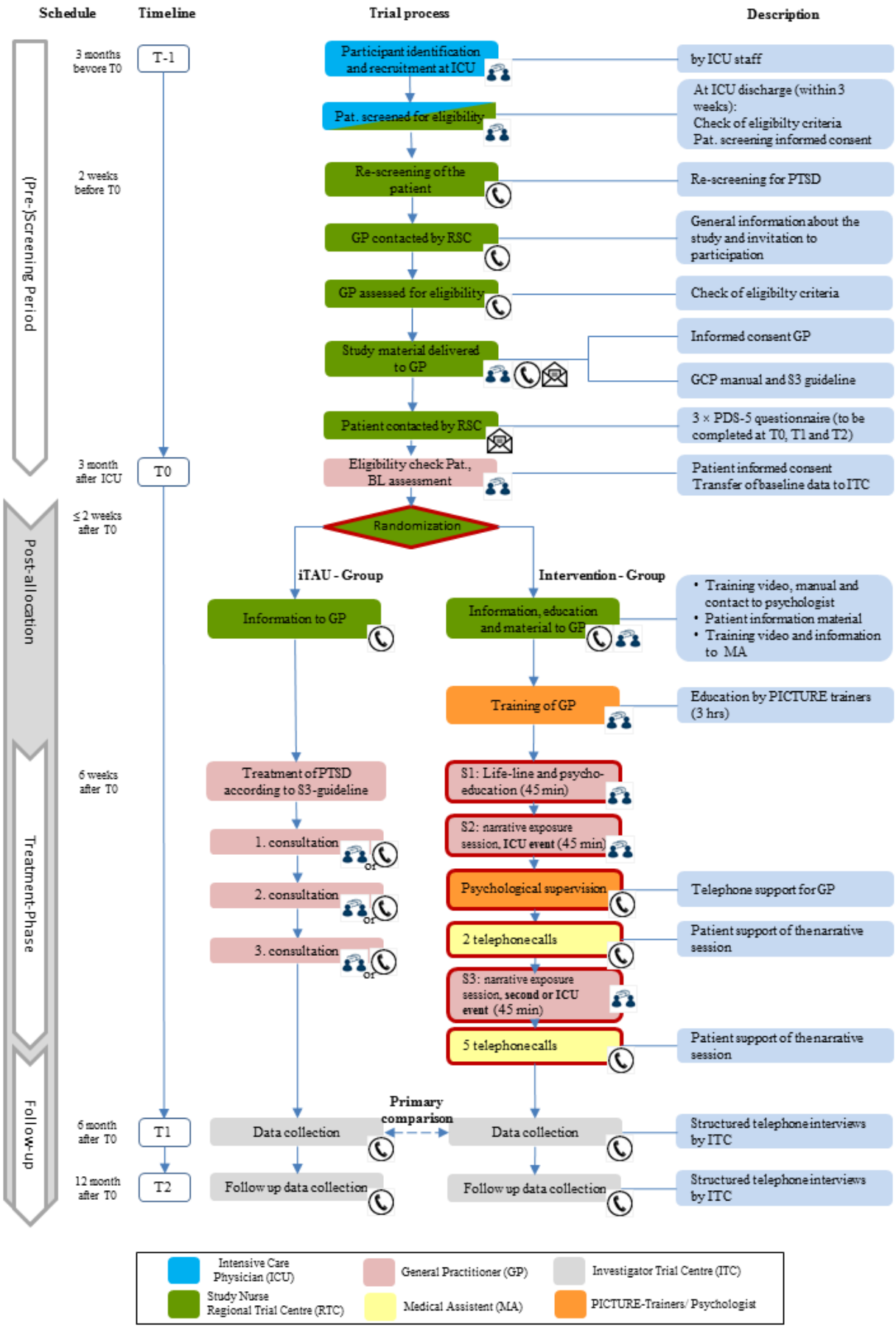

Figure 1*: Graphical depiction of study activities and components of intervention*

3 Trial Objectives and Endpoints

The overall goal of the PICTURE trial is to evaluate whether a primary care based NET-oriented intervention improves patient-reported outcomes such as PDS total severity score, quality of life, common co-morbidities depression and anxiety in patients with PTSD after ICU discharge. This study aims to describe and compare the real-world effectiveness safety and applicability of a primary care based complex psychological intervention with improved "usual care".

3.1 Add-on projects

Three sub-studies will be conducted together with the PICTURE – PTSD after ICU Survival-trial ("PICTURE-Economics", and "PICTURE-Psychotherapy in general practice").

3.2 Primary Objective

The primary objective of the trial is to determine the effect of the NET-oriented therapy compared to iTAU, on patient-reported PTSD symptoms measured by the PDS total severity score after 6 month.

Primary study hypothesis:

The NET–oriented intervention is more effective than iTAU in improvingPTSD symptoms measured via the PDS total severity score.

3.3 Primary Endpoint

The primary efficacy endpoint is defined as the absolute change in PDS-5 total severity score from baseline to T1, i.e. the difference between the 6-month post-randomization score and baseline score assessed at T0.

3.4 Secondary Objectives

The secondary objectives of this trial are:

a) To assess the effects of the NET-oriented intervention compared to iTAU with respect to

- *depression symptoms*
- *anxiety symptoms*

Hypothesis: Depression is a comorbid disorder commonly associated with PTSD (Campbell et al., 2007) and observed post ICU (Davydow, Gifford, Desai, Bienvenu, & Needham, 2009). Anxiety disorders are common prevalent comorbidities following posttraumatic symptoms (Zlotnick et al.,

2006). NET effects were demonstrated in reduction of severity/diagnosis, depression, suicidality, anxiety and drug abuse (Schauer, Neuner, & Elbert, 2011).

b) To assess the effects of the NET-oriented intervention compared to iTAU with respect to

- *health-related quality of life*
- *Disability*
- *Patient Activation Measure*

c) Objective of the health economic evaluation ("PICTURE-Economics")

To determine whether the NET intervention is cost-effective from a societal perspective compared to iTAU in patients diagnosed with PTSD after ICU survival.

We hypothesize that the NET-treated patients will cause lower health care costs and lower productivity losses compared to patients assigned to the iTAU condition while gaining a larger number of quality-adjusted life years, according to a better quality of life.

d) To check for the occurrence of adverse effects of the intervention

3.5 Secondary Endpoints

- PDS-5 total severity score: absolute change from baseline at T2
- PHQ-9 total score: absolute change from baseline to T1 and T2
- OASIS-D total score: absolute change from baseline to T1 and T2
- EQ-5D-5L: VAS at T1 and T2
- WHODAS 2.0 total score: absolute change from baseline to T1 and T2
- PAM total score: absolute change from baseline to T1 and T2

3.6 Additional secondary endpoints for "PICTURE-Economics"

- Costs (CSSRI)
- QALYs (calculated using the EQ-5D-5L index values)
- incremental cost-effectiveness ratio

at T1 and T2.

3.7 Safety Variables

The occurrence of safety events (SAE) between S1 in the NET-group/the first of three GP-consultations in the iTAU-group and T2, e.g. death, major depression or suicidality, hospitalization, and referral for psychiatric care, will be assessed systematically by the GP during the consultations. Any indications of SAE will be documented and reported to the PI and the data safety and management board (DSMB) (See section 13).

4 Trial Design

General design:
This investigator-initiated study is designed as a prospective, randomized, multi-center, two-arm parallel-group, assessor-blinded, controlled, comparative effectiveness trial with a fixed sample design.

4.1 Number of subjects

A total a number of N = 340 patients together with their treating GPs was calculated as the recruitment target. Recruitment was concluded on Dec. 31st 2022 with a total of N=318 randomized participants.

4.2 Time Schedule

Per patient:

- duration of intervention, NET group: 18 weeks (6 weeks narrative session plus 12 weeks telephone monitoring)
- duration of intervention, iTAU group: 3 consultations with the GP between randomization and T1
- duration of follow-up: 6 months
- total individual study duration: 12 months

Trial duration:

- Planned Start Date (Screening): 01.10.2017
- Planned Start Date – Enrollment (FPFV): 01.01.2018
- Planned End Date (LPLV): 31.12.2023 (including FUs)

The end of the clinical trial is defined by the last individual trial-specific examination during the last visit of the last patient to be part of the trial.

5 Trial Population, Eligibility Criteria, Recruitment

See study protocol.

6 Randomization and Blinding

Concealed randomization to both treatments (NET; iTAU) is performed with a 1:1 allocation ratio. The trial statistician will remain blinded to randomization codes throughout the course of the trial, i.e. until the study database has been finalised and locked for the final analyses after LPLV.

7 Trial Procedures

7.1 Methods of Assessment

The following section will give an overview and adequate explanations to the examinations and procedures (assessment instruments) that were performed in this trial. Screening instruments are described in section 8.7. Several clinical scores derived from patient questionnaires (mode of administration: paper-based self-administered or by telephone interview) were defined as primary and secondary outcomes.

7.1.1 PDS-5 total severity score (PTDS symptoms)

The PDS-5 total severity score is derived from the 'Posttraumatic Stress Diagnostic Scale', a patient-reported questionnaire assessing the PTSD-related symptoms according to DSM-5 [validated German translation, Wittmann et al. Eur J Psychotraumatol. 2021]. Each of the 22 items refers to symptoms experienced in the past month only and is answered on a 5-point Likert scale ranging from 0 (not at all) to 4 (more than 5 times per week/ severe). This results in a total score with a range from 0 to 88 points. The PDS-5 total severity score has been found to have excellent psychometric properties and correlates with similar instruments for the diagnosis of posttraumatic stress, with high scores serving more severe PTSD symptoms (Foa et al., 2016). In this trial, the PDS-5 total severity score serves as key inclusion and exclusion criterion for the severity of PTSD symptoms at the beginning of the study prior to randomization.

7.1.2 PHQ-9 (Depression)

Depressive symptoms experienced over the last 2 weeks will be assessed by means of the primary care validated Patient Health Questionnaire-9 (PHQ-9) (Kroenke & Spitzer, 2002). Each of the 9 items is scored from 0 'not at all' to 3 'nearly every day'. The PHQ-9 total sum score as a measure of depression severity ranges from 0 to 27, whereas a high score indicates severe impairment.

7.1.3 OASIS (Anxiety)

The brief OASIS questionnaire is the only measure of anxiety severity and impairment applicable to multiple anxiety disorders that has been validated for use in primary care (Campbell-Sills et al., 2009) (Norman et al., 2011). The five items are inspired by the ICD-10 F40-43 criteria (for phobic and other anxiety disorders, obsessive-compulsive disorder, reaction to severe stress and adjustment disorders) and refer to all aspects of anxiety symptoms, including panic attacks, situational anxieties, worries, flashbacks, hypervigilance of startle, experienced over the past week.
There are five different response options for each item, which are coded 0–4 and summed to obtain the OASIS total score ranging from 0 (no anxiety) to 20 points.The five items ask about anxiety and fear, including frequency, intensity of symptoms, avoidance behaviour and impairments in daily life through these symptoms.
The OASIS-D is the German version of the validated original Anglo-American questionnaire, which was translated according to international standards for cross-cultural adaptation of self-report measures (Beaton DE, 2000) (Hiller TS, 2014) (Zlotnick et al., 2006)

7.1.4 Disability (WHODAS 2.0 – short version)

The WHODAS 2.0 instrument is commonly used to assess disability (Üstün TB, Kostanjsek N, Chatterji S, & J, 2010). The WHODAS-2-short version contains 12 items on functioning and disability with a recall period of 30 days covering 6 domains: *Cognition* (2 items), *Mobility* (2 items), *Self-care* (2 items), *Getting along with others* (2 items), *Life activities* (2 items), and *Participation in society* (2 items). Response options go from 1 (no difficulty) to 5 (extreme difficulty or cannot do). WHODAS-2 scores are computed for each domain by adding the item responses and transforming them into a range from 0 to 100, with higher scores indicating higher levels of disability. A global sum-score across all domains are also computed. WHODAS 2.0 has good psychometric qualities, including good reliability and item-response characteristics, and its robust factor structure remains the same across cultures and in different patient populations (Üstün TB et al., 2010).

7.1.5 PAM-Score (Thirteen-Item Patient Activation Measure)

As the promotion of active participation is one of the duties of the GP (Hibbard, Mahoney, Stockard, & Tusler, 2005), (Hollnagel & Malterud, 1995), the PAM measures the active participation of patients and the self-management of their state of health (Hibbard, Stockard, Mahoney, & Tusler, 2004) in form of a self-assessment. A short German version (PAM13) has been developed for use in clinical practice and research, that has good reliability and validity (Brenk-Franz et al., 2015) (Zill et al., 2013). The 13-item self-administered questionnaire assesses the knowledge of the patient regarding his/her health problems, the ability and the confidence to cope with these problems independently. Each item has four response categories with scores from 1 to 4: (1) strongly disagree, (2) disagree, (3) agree, and (4) strongly agree. The fourth item has an additional category with (5) not applicable. Evaluation is made by adding the raw values with a range of 13-52. For standardization of the gross total value, the sum-scale will be calibrated to a 0 to 100 metric.

7.1.6 EQ-5D-5L (Quality of Life)

The EQ-5D-5L is a generic instrument to measure HRQoL, and validated for several modes of administration (in-person, phone, mail). It consists of five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) with five possible levels for each dimension, describing the severity of problems in the specific dimension experienced today (*EQ descriptive system*): no problems, slight problems, moderate problems, severe problems, extreme problems.

The EQ-5D-5L is applicable to a wide range of health conditions and treatments and provides a single index value derived from the severity of problems in the five dimensions. It takes only a few minutes to complete.

Additionally, the EQ-5D-5L includes a visual analogue scale (*EQ VAS*), a thermometer-like rating scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). Participants are asked to mark their current overall state of health on the scale.

7.1.7 Questionnaires for the health economic evaluation

The efficacy outcome of "PICTURE-Economics" is the *incremental cost-effectiveness ratio*. To assess the incremental cost-effectiveness ratio, two measures, one for costs, one for effects, are necessary.

Costs will be assessed by means of a modified (shortened) German version of the Client Sociographic and Service Receipt Inventory (**CSSRI**) (Chisholm et al., 2000). This questionnaire considers the resource utilization for inpatient services, outpatient physician services, outpatient therapeutic services, medications as well as formal and informal care. Additionally, the questionnaire captures productivity losses caused by absenteeism. Resource utilization will be monetarily valued by means of administrative and market prices according to the German manual for standardized unit costs by Bock et al (Bock et al., 2015). Absenteeism will be valued according to the human capital approach by means of gross hourly wage plus non-wage labor costs.

QUALYs*:* As measure of effects, we will use QALY, a composite measure consisting of the duration of life multiplied by a measure of preference-based HRQoL. In this study, the EQ-5D-5L index values will be employed (Herdman et al., 2011).

7.1.8 Covid-19 Burden Scale

The Covid-19 burden scale is a 10 item questionnaire adapted from Brailovska and Margraf 2020. It includes questions regarding the place of residence, elevated risk (elderly, immunosuppression, comorbidities), exposure to Covid-19, quarantine, feelings of fear or restriction, information and media usage, activities of daily life, healthcare utilization. This questionnaire was deployed in March 2020 at the beginning of the Covid-19 pandemic at T0, T1 and T2 interviews in order to account for effects due to the pandemic and lockdown measures.

7.2 Time schedule of Measurements

All visits and telephone calls will be performed according to the table 'Schedule of Activities and Assessments' (section 2) and the flowchart in Figure 1.

8 Statistical Methods

8.1 Planned Statistical Analyses

8.1.1 Analytical steps

1) Data cleaning and curation (completeness, plausibility, outliers, etc.)
2) Dropout analysis
3) Analysis of baseline data
   a) Descriptive analysis of all relevant clinical and sociodemographic characteristics, as well as outcome variables at baseline
   b) Cross-sectional exploratory analysis
4) Analysis of primary and secondary outcomes at primary endpoint T1
   a) Descriptive analysis, comparison of group differences
5) Longitudinal analysis
   a) Descriptive analysis of primary and secondary outcomes throughout observation period
   b) Calculation of multivariate statistical models from baseline to the primary endpoint T1 for primary and secondary outcomes
   c) Calculation of multivariate statistical models from baseline to the secondary endpoint T2 for primary and secondary outcomes

8.1.2 Analysis of the primary efficacy endpoint

Primary efficacy endpoint is the absolute change in the PDS total severity score from baseline at month 6 (Δ_6PDS := $PDS_{T1} - PDS_{T0}$). By default, the mode of administration is a self-administered paper-based version. For patients who did not complete and send back the paper-based patient questionnaire (non-responding survivors), the PDS-5 total score was assessed during the telephone survey T1, scheduled 6 months after randomization.

The null hypothesis H_0: $G_{NET}(x) = G_{iTAU}(x)$ and $K_{NET}(t) = K_{iTAU}(t)$ $(0 < t \leq T)$

is that the treatment groups NET and iTAU will not differ with respect to the distributions of the observed outcome measure Δ_6PDS, whereas $G_i(x)$ is the cumulative probability distribution of the observed change in PDS severity scores at T1 in group *i (i* = NET or iTAU*)*, and the distribution of times of death (most likely cause of missingness), $K_i(t)$ is the cumulative distribution of informative event times in group *i*.

The null hypothesis will be tested using a two-sample t-test. The null hypothesis can be rejected if the two-sided *p*-value related to the test statistic for the treatment effect is equal to or smaller than the significance level α=0.05 (two-sided).

Due to an expected high drop-out rate from death in critically ill patients, the nonparametric Lachin method was initially planned as the primary analysis, which is a modified version of the Wilcoxon-Mann-Whitney U-test with a worst rank approach for missing data. However, results from an interim report showed relatively low drop-out rates and a normal distribution of the main outcome variable (Δ_6PDS). We have consequently decided to update the SAP, replacing the originally planned nonparametric test with a regular two-sample t-test. This change is made to align with the observed normal distribution characteristics of the main outcome variable. The t-test is more suitable for normally distributed data and will provide more accurate estimates of the treatment effect compared to nonparametric tests.

The principal analysis will be performed according to the intention-to-treat (ITT) principle, and unadjusted for screening or baseline covariates or site. The significance level is set to alpha = 5% (two-sided).

Missings prior to the time of the follow-up measurement will occur because of an informative, disease-related event (e.g. death, morbidity) and for other reasons (e.g. non-responders at follow-up measurements T1, T2, loss to follow-up, consent withdrawn). To address the impact of several missingness mechanisms (MAR; MNAR) sensitivity analyses will be performed, e.g. by inverse probability weighting, mixed effect models assuming MAR ("all observed data approach") using the whole observed PDS profile of the surviving patient; multiple imputations techniques; or even complete case analyses using ANCOVA (absolute change score as response variable, treatment group as covariate, adjusting for the baseline score value) for responding survivors until T1.

Moreover, we plan sensitivity analyses in the per protocol population, using linear mixed effects models to explore the role of covariates where indicated (e.g. age and gender, GP-related factors, Covid-19 pandemic effects).

8.1.3 Analyses of secondary efficacy endpoints

All secondary efficacy analyses will be exploratory, i.e. performed without adjustment for multiplicity, using standard methods of inferential statistics appropriate for the given secondary outcome measure. Two-sided tests for detecting treatment differences will be carried out.

Descriptive comparisons of, e.g. change scores measured at T1 and T2, or patient characteristics, will be mainly conducted with the t-test or Mann-Whitney-Wilcoxon test or, in the case of a binary outcome, with the Fisher's exact test, as appropriate based on distribution characteristics. With respect to missing score values, the same considerations used for the primary outcome will apply equally to the pre-specified secondary efficacy outcomes.

Secondary outcomes will be used to explore sociodemographic and behavioral determinants of the study outcome throughout the observation period as well as their interaction with the intervention. Sub-studies will explore trajectories of secondary outcomes throughout the study period, as well as

the impact of the intervention on these outcomes, e.g. health-related quality of life, depression, healthcare utilization. Adjusted multivariate regression models (e.g. mixed effect models for longitudinal data, GLM for cross-sectional data) will be chosen based on the available data and distributions for further exploratory analyses of secondary outcomes. Qualitative studies will assess process indicators in order to evaluate applicability and feasibility of the proposed intervention.

Health economic evaluation

The cost-effectiveness of the NET intervention compared to iTAU will be determined from a societal perspective based on the ITT population. First, the incremental cost-effectiveness ratio (ICER) will be calculated as the difference in mean cost divided by the difference in mean QALYs:

ICER= (CostNET – CostsiTAU)/(QALYNET – QALYiTAU).

Second, net-benefit regressions will be conducted to determine the uncertainty of the point estimate (ICER) and to adjust for potential baseline differences and confounders (Hoch 2002). These results will be used to construct cost-effectiveness acceptability curves, which show the interventions' probability of being cost-effective at different willingness-to-pay margins (Range: 0€/QALY – 150.000€/QALY; raised in 10.000€/QALY steps) in comparison to iTAU. The underlying assumptions regarding the calculation of costs and benefits will be investigated in several sensitivity analyses.

PICTURE Psychotherapy

For a qualitative evaluation of the PICTURE intervention, semi-structured interviews were conducted in a small subset of NET-participants with patients and GPs after T1. An estimated 10 to 15 interviews of patients and GP's, respectively and 3-5 expert interviews of supervising psychologists is needed for qualitative analysis. The sampling procedure was done in a systematic way trying to involve a broad range of aspects. On enrolling interviewees we aimed to include patients and GP's from urban and rural background, male and female gender, different age groups, patients of different social background and GP's in single and partner practices and with varying length of working experience. Transcribed interview data will be analyzed using thematic analysis to explore process and implementation experiences (Braun & Clark, 2006). The analytic process is structured and includes 6 defined phases (Phase 1: Becoming familiar with the data, Phase 2: Generating initial codes, Phase 3: Searching for themes, Phase 4: Reviewing themes, Phase 5: Defining and naming themes, Phase 6: Producing the report).

8.1.4 Safety analysis

Safety analyses will be conducted for SAE reported during the trial period. The frequency of events and the possible relationship to the treatment condition will be analysed descriptively. The number and percentage of patients experiencing each SAE will be presented for each treatment arm. The number and percentage of occurrences of each SAE will also be presented for each treatment arm. No formal statistical testing will be undertaken. Data quality including potential underreporting will be discussed.

8.2 Interim analysis

A fixed sample design was planned without confirmatory statistical testing for early decision making, and will be conducted as planned. There is no pre-planned efficacy interim analysis.

8.3 Sample size calculation

We performed sample size calculations and additional simulations to detect a clinically relevant and empirically justified effect with respect to the PDS-5 total severity score (range 0-88 points). A detailed description of this process is given in the trial protocol. Based on these considerations, the achieved recruitment number of N=318 participants seems sufficient to satisfy the underlying assumptions of the sample calculations.

8.4 Definition of populations included in the analyses

This clinical trial will be analyzed according to the ITT principle. This means that the subjects will be analyzed in the treatment arms to which they were randomized, irrespective of whether they refused or discontinued the treatment, or whether other protocol violations are revealed.

The per-protocol (PP) population is a subset of the ITT population.

An analysis per-protocol (PP) will exclude or censor endpoint information considering major protocol deviations potentially effecting subjects' specific endpoint value, e.g., in the case of major violation of eligibility criteria (patient or GP level), lack of sufficient treatment per protocol (e.g., predefined number of sessions or consultations between T0 and T1) or unsatisfactory evaluations for endpoint assessment at predefined time points (telephone survey). The PP analyses excluding data from protocol non-adherers will be performed for the purpose of a sensitivity analysis and investigating robustness of results.

8.5 Protocol Violations

Protocol violations are major deviations from the procedures outlined in the study protocol. All protocol violations will be listed and the impact on the evaluation of the corresponding patient (or GP) will be discussed in a blinded manner prior to the statistical analyses.

8.6 Handling of Drop-outs, Withdrawal, and Missing Data

Subjects dropping out of the trial after randomization will be analyzed using all available data according to the ITT principle. Drop-outs will not be replaced.

Sensitivity analyses will include comparison of characteristics of drop-outs and participants, as well as comparison of adjusted and unadjusted statistical models.

9 Data Collection, Handling and Record Keeping

Data management was and will be performed at the Institute of General Practice at the site of the PI.

9.1 Data Forms and Data Entry

In the PICTURE trial, all data is documented on paper-based and electronic case report forms (eCRFs). This may be done at the ITC or at the RTC where the data are originated. Original study forms were entered in a web-based software tool (LibreClinica) and kept on paper file at the RTC.

9.2 Data Transmission and Editing

The data input screens are based on the CRFs. Data integrity is enforced by a variety of mechanisms, referential data rules, valid values, range checks, and consistency checks against data already stored in the database (i.e., longitudinal checks). Checks were applied at the time of data entry into a specific field and/or before the data is written (committed) to the database. Modifications to data written to the database were automatically documented through either the data change system or an inquiry system.

10 Reporting

10.1 Statistical Report

After completion and approval of the analyses by the responsible biostatistician, a statistical report will be prepared. Except when required by law, no one will disclose a result of the clinical trial to third parties unless all parties involved have first agreed on the results of the analyses and their interpretation.

11 Publication policy

The trial was prospectively registered in a public database (DRKS: http://www.germanctr.de; Clinical Trials: https://clinicaltrials.gov). Efficacy and safety results will be submitted for at least one main scientific publication in a peer-reviewed journal. Publication or lecture of data or trial results needs a previous annotation and approval of the PI. All subject-related data will be published in a pseudonymous form. The right of publication rests primarily with the PI and the other investigators and researchers involved.

All data collected in connection with the clinical trial will be treated in confidence by the PI and all others involved in the trial, until publication of the main results.

Interim data and final results may only be published (orally or in writing) with the agreement of the PI and the other investigators. This is indispensable for a full exchange of information between the above-named parties, which will ensure that the opinions of all parties involved have been heard

before publication. Details on authorship rules and publication strategy are provided in a separate document (PICTURE publication guidelines).

The agreement, which does not include any veto right or right of censorship for any of the parties involved, may not be refused without good reason.

Beaton DE, B. C., Guillemin F, Bosi Ferraz M. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *SPINE, 25*(24), 3186 - 3191. Retrieved from https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0cHVod2J8Z3g6MmRkNzk4N2EwOGZmZGNjMw

Bock, J. O., Brettschneider, C., Seidl, H., Bowles, D., Holle, R., Greiner, W., & Konig, H. H. (2015). [Calculation of standardised unit costs from a societal perspective for health economic evaluation]. *Gesundheitswesen, 77*(1), 53-61. doi:10.1055/s-0034-1374621

Brenk-Franz, K., Strauss, B., Tiesler, F., Fleischhauer, C., Ciechanowski, P., Schneider, N., & Gensichen, J. (2015). The Influence of Adult Attachment on Patient Self-Management in Primary Care--The Need for a Personalized Approach and Patient-Centred Care. *PLoS One, 10*(9), e0136723. doi:10.1371/journal.pone.0136723

Campbell, N. C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., . . . Kinmonth, A. L. (2007). Designing and evaluating complex interventions to improve health care. *BMJ, 334*(7591), 455-459. doi:10.1136/bmj.39108.379965.BE

Campbell-Sills, L., Norman, S. B., Craske, M. G., Sullivan, G., Lang, A. J., Chavira, D. A., . . . Stein, M. B. (2009). Validation of a brief measure of anxiety-related severity and impairment: the Overall Anxiety Severity and Impairment Scale (OASIS). *J Affect Disord, 112*(1-3), 92-101. doi:10.1016/j.jad.2008.03.014

Chisholm, D., Knapp, M. R., Knudsen, H. C., Amaddeo, F., Gaite, L., & van Wijngaarden, B. (2000). Client Socio-Demographic and Service Receipt Inventory--European Version: development of an instrument for international research. EPSILON Study 5. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. *Br J Psychiatry Suppl*(39), s28-33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10945075

Davydow, D. S., Gifford, J. M., Desai, S. V., Bienvenu, O. J., & Needham, D. M. (2009). Depression in general intensive care unit survivors: a systematic review. *Intensive Care Med, 35*(5), 796-809. doi:10.1007/s00134-009-1396-5

Foa, E. B., McLean, C. P., Zang, Y., Zhong, J., Powers, M. B., Kauffman, B. Y., . . . Knowles, K. (2016). Psychometric properties of the Posttraumatic Diagnostic Scale for DSM–5 (PDS–5). *Psychological Assessment, 28*(10), 1166-1171. doi:10.1037/pas0000258

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., . . . Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res, 20*(10), 1727-1736. doi:10.1007/s11136-011-9903-x

Hibbard, J. H., Mahoney, E. R., Stockard, J., & Tusler, M. (2005). Development and testing of a short form of the patient activation measure. *Health Serv Res, 40*(6 Pt 1), 1918-1930. doi:10.1111/j.1475-6773.2005.00438.x

Hibbard, J. H., Stockard, J., Mahoney, E. R., & Tusler, M. (2004). Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Serv Res, 39*(4 Pt 1), 1005-1026. doi:10.1111/j.1475-6773.2004.00269.x

Hiller TS, B. J., Brenk-Franz K, Storch M, Schneider N, Schelle M, Thiel P, Gensichen J. (2014). *Evaluation der deutschen Version der Overall Anxiety Severity and Impairment Scale (OASIS-D) für den Einsatz in der Hasuarztpraxis*. Paper presented at the 48. Kongress für

Allgemeinmedizin und Familienmedizin, Hamburg. Vortrag retrieved from http://www.allgemeinmedizin.uni-jena.de/content/forschung/instrumente/angststoerungen/index_ger.html

Hollnagel, H., & Malterud, K. (1995). Shifting attention from objective risk factors to patients' self-assessed health resources: a clinical model for general practice. *Fam Pract, 12*(4), 423-429. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8826059

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr Ann, 32*(9), 509-515.

Needham, D. M., Davidson, J., Cohen, H., Hopkins, R. O., Weinert, C., Wunsch, H., . . . Harvey, M. A. (2012). Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. *Crit Care Med, 40*(2), 502-509. doi:10.1097/CCM.0b013e318232da75

Norman, S. B., Campbell-Sills, L., Hitchcock, C. A., Sullivan, S., Rochlin, A., Wilkins, K. C., & Stein, M. B. (2011). Psychometrics of a brief measure of anxiety to detect severity and impairment: the Overall Anxiety Severity and Impairment Scale (OASIS). *J Psychiatr Res, 45*(2), 262-268. doi:10.1016/j.jpsychires.2010.06.011

Schauer, M., Neuner, F., & Elbert, T. (2011). *Narrative Exposure Therapy: A Short-Term Treatment for Traumatic Stress Disorders, 2nd revised and expanded edition*. Cambridge, MA, USA; Göttingen, Germany: Hogrefe.

Üstün TB, Kostanjsek N, Chatterji S, & J, R. (2010). *Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0)*. Geneva: WHO

Zill, J. M., Dwinger, S., Kriston, L., Rohenkohl, A., Härter, M., & Dirmaier, J. (2013). Psychometric evaluation of the German version of the patient activation measure (PAM13). *BMC Public Health, 13*(1), 1027. doi:10.1186/1471-2458-13-1027

Zlotnick, C., Johnson, J., Kohn, R., Vicente, B., Rioseco, P., & Saldivia, S. (2006). Epidemiology of trauma, post-traumatic stress disorder (PTSD) and co-morbid disorders in Chile. *Psychol Med, 36*(11), 1523-1533. doi:10.1017/S0033291706008282