Predictors and trajectories of treatment response to SSRIs in patients suffering from PTSD

Nøhr, Anne Krogh; Eriksson, Hans; Hobart, Mary; Moltke, Ida; Buller, Raimund; Albrechtsen, Anders; Lindgreen, Stinus

Published in:
Psychiatry Research

DOI:
10.1016/j.psychres.2021.113964

Publication date:
2021

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Predictors and trajectories of treatment response to SSRIs in patients suffering from PTSD

Anne Krogh Nøhr, Hans Eriksson, Mary Hobart, Ida Moltke, Raimund Buller, Anders Albrechtsen, Stinus Lindgreen

The Bioinformatics Centre, Department of Biology, University of Copenhagen, Copenhagen N, Denmark
H. Lundbeck A/S, Valby, Copenhagen, Denmark
Osaka Pharmaceutical Development & Commercialization Inc., 508 Carnegie Center Drive, Princeton, NJ 08540, USA

1. Background

Posttraumatic stress disorder (PTSD) is a common mental disorder with an estimated lifetime prevalence of 8-12% in the general population (Kessler, 2000, 1995). PTSD develops in response to a traumatic event and the symptoms include re-experiencing the trauma, avoiding reminders of the trauma, negative alterations in cognition and mood, and hyperarousal (Pai et al., 2017). Additionally, patients suffering from PTSD often suffer from one or more comorbidities, among which the most common are depression, anxiety, and substance use disorder (Brady et al., 2000).

Treatment of PTSD consists of psychotherapy, pharmacotherapy or a combination of these and the treatment aims to reduce severity of symptoms, treat comorbid disorders, prevent relapse and improve quality of life (Forbes et al., 2010; Ursano et al., 2004). The selective serotonin reuptake inhibitors (SSRIs) sertraline and paroxetine are the recommended first-line medications in most treatment guidelines and the only drugs for PTSD approved by the US Food and Drug Administration (FDA) (Forbes et al., 2010). Even though both medicines have shown better results compared to placebo (Hoskins et al., 2015; Watts et al., 2013), they only lead to a response rate of 37-62% in patients suffering from PTSD (Brady et al., 2000; Davidson et al., 2001; Marshall et al., 2001; Stein et al., 2003). In these studies response to treatment was defined as a >30% reduction from baseline on the Clinician Administered PTSD Scale (CAPS) score (Blake et al., 1990) and a rating of “much” or “very much” improvement on the Clinical Global Impressions scale Improvement (CGI-I) (Guy, 1976).

Given the limited response rate, it is clear that SSRI treatment is not effective for all patients. A better understanding is needed of the factors which influence the individual differences in treatment response in order to predict the optimal treatment regime for patients. Furthermore, identification of latent patient classes based on SSRI treatment response, like fast responders, can potentially be used for designing clinical trials. To move into the field of precision medicine, knowledge about predictors of treatment response is crucial, and some of these individual differences can be explained by heterogeneity in response trajectories.

ARTICLE INFO
Keywords:
Posttraumatic stress disorder
Serotonin reuptake inhibitors
Latent classes
Growth mixture modeling
Time since index trauma
Depression
Anxiety

ABSTRACT
Paroxetine and sertraline are the only FDA approved drugs for treatment of posttraumatic stress disorder (PTSD). Although both drugs show better outcomes than placebo, not all patients benefit from treatment. We examined predictors and latent classes of SSRI treatment response in patients with PTSD. Symptom severity was measured over a 12-week period in 390 patients suffering from PTSD treated with open-label sertraline or paroxetine and a double-blinded placebo. First, growth curve modeling (GCM) was used to examine population-level predictors of treatment response. Second, growth mixture modeling (GMM) was used to group patients into latent classes based on their treatment response trajectories over time and to investigate predictors of latent class membership. Gender, childhood sexual trauma, and sexual assault as index trauma moderated the population-level treatment response using GCM. GMM identified three classes: fast responders, responders with low pretreatment symptom severity and responders with high pretreatment symptom severity. Class membership was predicted based on time since index trauma, severity of depression, and severity of anxiety. The study shows that higher severity of comorbid disorders does not result in an inferior response to treatment and suggests that patients with longer time since index trauma might particularly benefit from treatment with sertraline or paroxetine.

Received 2 March 2020; Accepted 21 April 2021
Available online 26 April 2021
0165-1781/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
A.K. Nohr et al.

Psychiatry Research 301 (2021) 113964

2

and confirmed by the Mini International Neuropsychiatric Interview (MINI). The patients also needed a Clinically Administered PTSD Scale (CAPS) Part2 total score ≥ 70 at screening and baseline. Key exclusion criteria included: index trauma happened more than 15 years before the screening, significant risk of suicidality (verified as answer “yes” to the ideation items 4 and 5, or answer “yes” to any of the 5 behavior items on Columbia Suicide Severity Rating Scale), major depressive episode, anxiety disorder, substance abuse and alcohol dependence within 6 months prior to screening. The study was conducted at 59 sites located in 11 different countries: USA, Poland, Finland, South Africa, Serbia, Sweden, Mexico, Estonia, Argentina, France, and Italy. To ensure our findings were not due to differences between countries it was tested if adjustment of county changed the outcome of the analysis. The study was approved by the respective Competent Authorities and IRBs/Ethics Committees.

2.2. Measures

Visits were scheduled for week 1, 2, 4, 8, and 12. At each visit symptom severity of PTSD was assessed using CAPS Part2. The CAPS Part 2 contains questions about persistent re-experience, persistent avoidance of stimuli and persistent symptoms of increased arousal. The scale contains 17 items, each rated on a five point scale for frequency (0 is never or none and 4 is daily or almost every day) and intensity (0 is none and 4 is extreme). The 17-item scale gives a total score ranging from 0 to 136. The outcome measure was CAPS Part2 total score.

Candidate predictors were measured at baseline. Depressive symptom severity and anxiety symptom severity were assessed using the hospital anxiety and depression scale (HADS) (Zigmond and Snaith, 1983). HADS is designed to screen for anxiety and depression using two sub-scales; the A-scale measures anxiety and the D-scale measures depression. Both sub-scales have seven items, each ranging from 0 (absent) to 3 (maximum severity), leaving the total score of both scales to range from 0 to 21. The Life Events Checklist (LEC) (Gray et al., 2004) was used to assess the index trauma and contains 17 groups of traumatic events. Due to the large number of groups, the index traumas were combined into 6 overall groups. Data about time since index trauma, childhood trauma, demographic characteristics (age and gender) were also collected.

2.3. Data analysis

The primary methods used for the data analysis were growth curve modelling (GCM) and growth mixture modelling (GMM). The R package lcmm (Proust-Lima et al., 2017), version 1.7.9, was used for all GCM and GMM analysis. The method in the lcmm package is based on maximum likelihood and uses a modified Marquardt algorithm for optimization (Proust-Lima et al., 2017). Participants with at least one completed post-baseline assessment session (N = 390) were included in the analysis. We use these methods instead of standard use of ANCOVA and Generalized Linear Models (GLM) because they allow us to explore more flexible disease trajectories where change is not linear. In addition, we can model latent structures of the data using these methods to identify groups of individuals with different disease trajectories.

In the first part of the analysis, GCM was used to explore the mean treatment response trajectory and how it was affected by predictors. Briefly, GCM is a method developed to describe a population’s mean trajectory in time and the mean trajectory is modelled to follow a specific curve (e.g. linear, quadratic). Likewise, a trajectory is modelled for each individual in the population, assuming to follow a curve of the same type with curve parameters that are specific to the individual patient. This makes the method robust to missing data, as the missing data is estimated from the individual patient’s trajectory. In addition, GCM takes advantage of correlations among repeated measurements. More specifically, treatment response, $y_i$, for participant $i$ was modelled across the 5 visits using the following equation

$$y_i = X_i\beta + Z_i\alpha + \epsilon_i$$

(1)
where \( \beta \) is the vector of fixed effects (population level regression coefficients) of length \( p \) associated with the design matrix \( X_i \) of size \( 5 \times p \), and \( u_i \) is a vector of random effects (subject specific regression coefficients) with length \( q \) associated with the design matrix \( Z_i \) of size \( 5 \times q \). The random effects were modelled to have zero mean with Gaussian random deviations and an unstructured variance-covariance matrix. \( e_i \) is the error vector of length 5 and the errors were independent and Gaussian with variance \( \sigma^2 \). Treatment response was modelled in three different ways: linear, linear and quadratic, and piecewise linear. Piecewise linear models were made because they can capture rapid changes; \( Z_i \) and \( X_i \) can capture both the linear and quadratic shapes of trajectories. The piecewise linear model was made by specifying a separate linear function between baseline and visit 1, 2, 3, 4 and 5. The predictors’ effects were modelled constant across time by including them as fixed effects in the model.

In the second part of the analysis, GMM was used to investigate heterogeneity of treatment response to define sub-populations; classes of patients with similar response trajectories. It was also explored if any predictors affected the class the patients belonged to. The main difference between the GMM and GCM is that participants are assumed to belong to one of \( G \) latent classes based on similarity of their treatment response trajectories. Each of the \( G \) latent classes has its own GCM, and thus a mean trajectory for each class and a trajectory for each patient is modelled. The only difference with the GCM in Eq. 1 is that the model includes class-specific fixed effects and class-specific distributions of the individual random effects. The treatment response trajectory for class \( g \) (\( g=1,2,...,G \)) can be expressed as:

\[
y_i | c_i = g = W_i v_i + Z_i u_i + e_i
\]

where \( c_i \) is a discrete variable that defines what latent class a participant belongs to. \( v_i \) is the class-specific fixed effects with design matrix \( W_i \), and \( u_i \) is still the individual random effects but now with class-specific distributions.

Unconditional two- through five-class linear, linear and quadratic, and piecewise linear models were made (the one-class GMM corresponds to GCM). Grid search was used to find the best set of initial values for the unknown growth parameters, resulting in the best possible models. Several fit indices, including Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC) and Bootstrap Likelihood Ratio Test (BLRT) were used to pick the model with the number of classes that best fitted the data. After identifying the best unconditional model, the effects of the predictors on the class specific growth parameters as well as class membership were explored. The impact of predictors on the class specific growth parameters were investigated by extending the conditional model to include predictors in the class-specific fixed effects of Eq. 2. The association of covariates with probability of class membership were expressed using a multinomial logistic specific fixed effects of Eq. 2. The association of covariates with probability of class membership were explored. The impact of predictors affected the class the patients belonged to. The main difference between the GMM and GCM is that participants are assumed to experience ‘GCM). Grid search was used to find the best set of initial values for the unknown growth parameters, resulting in the best possible models. Several fit indices, including Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC) and Bootstrap Likelihood Ratio Test (BLRT) were used to pick the model with the number of classes that best fitted the data. After identifying the best unconditional model, the effects of the predictors on the class specific growth parameters as well as class membership were explored. The impact of predictors on the class specific growth parameters were investigated by extending the conditional model to include predictors in the class-specific fixed effects of Eq. 2. The association of covariates with probability of class membership were expressed using a multinomial logistic specific fixed effects of Eq. 2. The association of covariates with probability of class membership were explored. The impact of predictors affected the class the patients belonged to. The main difference between the GMM and GCM is that participants are assumed to include in the analysis. Out of the remaining 390 participants, 23 withdrew before visit 2, 37 before visit 3, 54 before visit 4, 36 before visit 5. Fisher’s exact test and t-tests were used to investigate demographics and other baseline characteristics for participants that withdrew. A significant difference was found for age (t(388)=2.55, \( p=0.01 \), equal variance between the two groups are assumed). The participants who withdrew (mean=38.86 years, SD=11.50) were younger than those who completed the 12-week prospective treatment period (mean=41.91 years, SD=11.51). The overall response rate was 228 out of the 390 patients (58.5 %), with response defined as a >30% reduction from baseline on the total CAPS score.

### 3.2. Growth curve modelling

The unconditional linear, linear and quadratic, and piecewise linear GCM models all showed a linear decrease in the CAPS score per visit. Consequently, the simplest model, the linear model, was chosen as the best GCM. The linear GCM suggested that the average CAPS score was shown nominal significant univariate moderation effects on the slope (see Table 2): gender, childhood sexual abuse and sexual assault as index trauma. Females responded better compared to males. Patients who had experienced sexual assault as a child or had sexual assault as index

### 3. Results

#### 3.1. Sample characteristics

Table 1 presents the demographics, baseline characteristics and reasons for dropping out. At baseline, 407 participants were assessed, of which 17 (Protocol Violation 2, withdrawal of consent 2, adverse event 8, other 5) withdrew before the first assessment session and are not included in the analysis. Out of the remaining 390 participants, 23 withdrew before visit 2, 37 before visit 3, 54 before visit 4, 36 before visit 5. Fisher’s exact test and t-tests were used to investigate demographics and other baseline characteristics for participants that withdrew. A significant difference was found for age (t(388)=2.55, \( p=0.01 \), equal variance between the two groups are assumed). The participants who withdrew (mean=38.86 years, SD=11.50) were younger than those who completed the 12-week prospective treatment period (mean=41.91 years, SD=11.51). The overall response rate was 228 out of the 390 patients (58.5 %), with response defined as a >30% reduction from baseline on the total CAPS score.

#### 3.2. Growth curve modelling

The unconditional linear, linear and quadratic, and piecewise linear GCM models all showed a linear decrease in the CAPS score per visit. Consequently, the simplest model, the linear model, was chosen as the best GCM. The linear GCM suggested that the average CAPS score was shown nominal significant univariate moderation effects on the slope (see Table 2): gender, childhood sexual abuse and sexual assault as index trauma. Females responded better compared to males. Patients who had experienced sexual assault as a child or had sexual assault as index

### 3. Results

#### 3.1. Sample characteristics

Table 1 presents the demographics, baseline characteristics and reasons for dropping out. At baseline, 407 participants were assessed, of which 17 (Protocol Violation 2, withdrawal of consent 2, adverse event 8, other 5) withdrew before the first assessment session and are not included in the analysis. Out of the remaining 390 participants, 23 withdrew before visit 2, 37 before visit 3, 54 before visit 4, 36 before visit 5. Fisher’s exact test and t-tests were used to investigate demographics and other baseline characteristics for participants that withdrew. A significant difference was found for age (t(388)=2.55, \( p=0.01 \), equal variance between the two groups are assumed). The participants who withdrew (mean=38.86 years, SD=11.50) were younger than those who completed the 12-week prospective treatment period (mean=41.91 years, SD=11.51). The overall response rate was 228 out of the 390 patients (58.5 %), with response defined as a >30% reduction from baseline on the total CAPS score.
piecewise linear GMMs were made to find the model that best fitted the
difference in treatment response between paroxetine and sertraline.

Fit indices for growth mixture models.

covariate age also showed a significant moderation effect.
sault as index trauma still had a significant univariate moderation effect
was also conducted with country as a covariate in the models (see
alized in the Supplementary Material Figure 4. There was no significant
in the growth curve model.

Overview of how the individual candidate predictor affects slope and intercept
in the growth curve model.

Table 2

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Intercept (baseline CAPS)</th>
<th>Slope (CAPS/Visit)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect size (SE)</td>
<td>p-value</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (male)</td>
<td>-1.42 (1.55)</td>
<td>0.36</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.05 (0.06)</td>
<td>0.42</td>
</tr>
<tr>
<td>Index trauma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual assault (yes)</td>
<td>1.52 (2.24)</td>
<td>0.50</td>
</tr>
<tr>
<td>Combat or exposure to a war-zone (yes)</td>
<td>-0.64 (2.49)</td>
<td>0.80</td>
</tr>
<tr>
<td>Trauma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time since index trauma (years)</td>
<td>-0.43 (0.17)</td>
<td>0.01</td>
</tr>
<tr>
<td>Childhood sexual trauma (yes)</td>
<td>0.42 (1.93)</td>
<td>0.83</td>
</tr>
<tr>
<td>Scales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADS Anxiety</td>
<td>1.27 (0.18)</td>
<td>0</td>
</tr>
<tr>
<td>HADS Depression</td>
<td>1.26 (0.19)</td>
<td>0</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSRI (Paroxetine)</td>
<td>-0.95 (1.50)</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Note: AIC = Akaike’s Information Criterion; BIC = Bayesian Information Criterion; BLRT = Bootstrap Likelihood Ratio Test (1000 bootstraps). The best fitted model is in italics.

trauma responded better than those who had not. Sexual assault as index trauma is the predictor with the largest effect on the slope and is visualized in the Supplementary Material Figure 4. There was no significant difference in treatment response between paroxetine and sertraline.

The study was conducted in 11 different countries, which is a possible confounding variable. Therefore, the analysis in this section was also conducted with country as a covariate in the models (see Supplementary Material Table 6). Including country as a covariate showed similar results. Gender, childhood sexual abuse and sexual assault as index trauma still had a significant univariate moderation effect on the slope. The only difference was, when country was added as a covariate age also showed a significant moderation effect.

3.3. Growth mixture modelling - unconditional model

Unconditional 2- through 6-class linear, quadratic and piecewise linear GMMs were made to find the model that best fitted the data. The fit indices are shown in Table 3. By visual inspection the piecewise linear models were a better fit since the linear and quadratic models could not capture the fast response to treatment observed for some of the patients. The 3-class piecewise linear model was picked as the best fitted model since the model had the second lowest BIC and a significant BLRT. Thus, treatment response in this study can be best described as a mixture of three latent classes. The three classes will be referred to as “Responders with low symptom severity,” “Fast responders,” and “Responders with high symptom severity”. The three classes are further explored and described in the next section.

3.4. Growth mixture modelling - conditional model

Each predictor was entered into the model to test if they individually influenced what class the patients belonged to, see Table 4. Pretreatment severity of depression, pretreatment severity of anxiety and time since index trauma significantly predicted class membership. Subsequently, backward variable selection, with BIC as goodness-of-fit criterion, was used to find the optimal combination of predictors. The model with all three predictors was the best model. Fig. 1 shows the 3-class piecewise linear model with severity of depression, severity of anxiety, and time since index trauma as predictors of class membership.

Based on the maximum posterior probability (MAP) individuals were classified into the 3 classes. The largest class was the “Responders with low symptom severity” and it included 260 patients (67.01%), that were classified with a mean posterior probability of 92%. This class had the lowest CAPS score at baseline (84.52, SE = 0.69) and the score steadily declined over time. The “Fast responders” class comprised 63 patients (16.24%) and had a mean posterior probability for classification of 94%. Patients in this class had a CAPS score at baseline of 88.13 (SE = 1.25) and their decrease in symptom severity from baseline to the first week was steep after which the decrease in symptom severity flattened. The “Responders with high symptom severity” class contained 65 patients (16.75%) classified with a mean posterior probability of 87%. This class had the lowest CAPS score at baseline (84.52, SE = 0.69) and the symptom severity steadily declined over time.

Table 5 gives an overview of the 3-class piecewise linear model with pretreatment severity of depression, pretreatment severity of anxiety and time since index trauma as predictors of class membership.

Based on the maximum posterior probability (MAP) individuals were classified into the 3 classes. The largest class was the “Responders with low symptom severity” and it included 260 patients (67.01%), that were classified with a mean posterior probability of 92%. This class had the lowest CAPS score at baseline (84.52, SE = 0.69) and the score steadily declined over time. The “Fast responders” class comprised 63 patients (16.24%) and had a mean posterior probability for classification of 94%. Patients in this class had a CAPS score at baseline of 88.13 (SE = 1.25) and their decrease in symptom severity from baseline to the first week was steep after which the decrease in symptom severity flattened. The “Responders with high symptom severity” class contained 65 patients (16.75%) classified with a mean posterior probability of 87%. This class had the most severe pretreatment CAPS score of 106.29 (SE = 1.63), and the symptom severity steadily declined over time.

Table 5 gives an overview of the 3-class piecewise linear model with pretreatment severity of depression, pretreatment severity of anxiety and time since index trauma as predictors of class membership using “Responders with low symptom severity” as reference class. The

Table 4

Overview of how the individual predictors effect class membership when included in the 3-class growth mixture model. “Responders with low symptom severity” is the reference class.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>LRT (p-value)</th>
<th>Fast responders Effect size (SE)</th>
<th>Responders with high symptom severity Effect size (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (male)</td>
<td>0.51</td>
<td>-0.30 (0.33)</td>
<td>0.34 (0.38)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.15</td>
<td>-0.03 (0.01)</td>
<td>-0.004 (0.01)</td>
</tr>
<tr>
<td>Index trauma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual assault (yes)</td>
<td>0.18</td>
<td>0.79 (0.42)</td>
<td>0.06 (0.55)</td>
</tr>
<tr>
<td>Combat or exposure to a war-zone (yes)</td>
<td>0.65</td>
<td>0.43 (0.45)</td>
<td>0.11 (0.54)</td>
</tr>
<tr>
<td>Trauma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time since index trauma (years)</td>
<td>0</td>
<td>0.13 (0.03)</td>
<td>-0.04 (0.04)</td>
</tr>
<tr>
<td>Childhood sexual trauma (yes)</td>
<td>0.35</td>
<td>0.22 (0.40)</td>
<td>0.59 (0.41)</td>
</tr>
<tr>
<td>Scales</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADS Anxiety</td>
<td>0</td>
<td>0.01 (0.04)</td>
<td>0.40 (0.07)</td>
</tr>
<tr>
<td>HADS Depression</td>
<td>0</td>
<td>-0.01 (0.04)</td>
<td>0.31 (0.06)</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSRI (Paroxetine)</td>
<td>0.76</td>
<td>-0.04 (0.33)</td>
<td>-0.26 (0.35)</td>
</tr>
</tbody>
</table>

Note: LRT = Likelihood Ratio Test; SE = Standard Error; HADS = Hospital Anxiety and Depression Scale.
Responders with high symptom severity were characterized by a larger risk of high pretreatment depression and anxiety symptom severity. Conversely, low depression and anxiety symptom severity increased the patients likelihood of belonging to the “Fast responders” class or the “Responders with low symptom severity” class. Time since index trauma was the only predictor showing a significant difference between the classes “Fast responders” and “Responders with low symptom severity”. A shorter time since index trauma increased the possibility that a patient belonged to “Responders with low symptom severity” and a longer time since index trauma increased the likelihood of class membership in “Fast responders”. Time since index trauma was the only predictor that significantly separated “Responders with low symptom severity” from “Fast responders” and its effect on class membership for each class is visualized in Fig. 2.

4. Discussion

This study investigated trajectories and predictors of treatment response to sertraline and paroxetine for patients suffering from PTSD. In the first part of the analysis, GCM showed that gender, sexual assault as index trauma and childhood sexual trauma affected how well patients responded to treatment, when they were studied as a homogeneous population. In the second part of the analysis, the heterogeneity of treatment response was investigated using GMM and it was found that the patients’ response trajectories were best described by three latent classes: “Fast responders”, “Responders with low symptom severity”, and “Responders with high symptom severity”. In addition, pretreatment severity of depression, pretreatment severity of anxiety and time since index trauma predicted which class the individual patient was most likely to belong to.

Both severity of depression (Elliott et al., 2005; Phelps et al., 2017; Schumm et al., 2013; Stein et al., 2012) and severity of anxiety (Allan et al., 2016; Elliott et al., 2005) have previously been shown to predict class membership for other treatment types of PTSD using GMMS. In the current study and in previous studies (Phelps et al., 2017; Schumm et al., 2013; Stein et al., 2012), classes with high pretreatment PTSD symptom severity were also the classes with high severity of depression and vice versa. The same pattern is observed for severity of anxiety in the current study. In the GCM model, depression and anxiety symptoms did not predict the population level response (slope) but did predict the symptom severity of PTSD at baseline (intercept). Patients with higher depression and anxiety symptoms at baseline also had higher PTSD symptom severity. Hence it is likely that the role of depression and anxiety symptoms as predictors of class membership can be explained by the correlation between the severity of the comorbid conditions and the PTSD severity at baseline. Phelps et al. (2017) argued that a combination of several comorbidities (guilt and severity of depression and PTSD) impede treatment response more than an increased level of one comorbidity alone. Our study shows that patients with high PTSD symptoms at baseline are more likely to have high depression and anxiety symptoms. However, both the GCM and the GMM in this study, indicate that high depression and anxiety symptoms do not result in an inferior sertraline or paroxetine treatment response. Patients with high depression and anxiety symptoms respond just as well or better compared to
patients with low comorbid symptoms. It is important to note that patients diagnosed with comorbid depression and/or anxiety were not included in the study. Therefore, this study draws conclusions about the association between PTSD and the symptom severity of depression and anxiety only in patients where the diagnostic criteria for these comorbidities were not met.

Time since index trauma predicted both pretreatment severity of PTSD (intercept) in the GCM and class membership in the GMM. The GCM indicated that longer time since index trauma was associated with a lower pretreatment severity of PTSD at the population level. In the GMM, time since index trauma was an important predictor for the "Fast responders" and the "Responders with low symptom severity" even though their pretreatment severity of PTSD (intercept) was close to each other. Longer time since index trauma increased the patient’s likelihood of belonging to the "Fast responders" whereas the chance of belonging to the "Responders with low symptom severity" decreased. These findings are in contradiction with the literature, where longer time since index trauma is associated with less treatment response (Ehlers et al., 2013).

Gender, childhood sexual trauma and sexual assault as index trauma all individually affected the population level treatment response in the GCM, but none of them predicted class membership in the GMM. This suggests that they do not predict a specific response pattern, but that they have a population level effect independent of class membership. Childhood abuse is a known factor associated with poor response during treatment for depression (Miniati et al., 2010; Nemeroff et al., 2003). In this study patients that had experienced childhood sexual trauma responded better than those that had not, which contradicts prior results. Type of index trauma is also suggested to be a potential predictor of treatment response, since war veterans have shown a smaller effect (Watts et al., 2013). In this study, patients with combat related trauma did not show a separate effect on treatment response, but patients with an index trauma related to sexual assault responded better than patients with other index trauma. However, it is worth noting that although patients with combat related trauma do not respond as well as patients with trauma related to sexual assault they still show a treatment response to some degree (note only 10.3 % of the patients have combat related trauma; equivalent to 40 patients). This is interesting as there are no positive studies with sertraline versus placebo or paroxetine versus placebo for this patient group to our knowledge (Hoskins et al., 2015). The reason why this group responds poorly is not yet understood. Consistent with the findings from the meta-analysis by Watts et al. (2013), we saw in the GCM that women tended to respond better than men. This could explain why patients with combat related trauma do not show a larger treatment response since the majority is men (92.5 %). This could also explain why childhood sexual trauma and sexual trauma as index trauma appeared to result in larger treatment responses, since the majority in these groups were women (90.4 % of patients with childhood sexual trauma and 92 % of patients with sexual trauma as index trauma).

4.1. Limitations and future directions

The current study was prospective and conducted at 59 sites in 11 different countries. Country is a possible confounding variable and was accounted for by running the first part of the analysis with country as a covariate in the GCMs. Including the country as a covariate did not change the outcome of this analysis. It was not tested for the confounding effect of country in the second part of the analysis since the inclusion of covariates in the GMMs would result in models with too many parameters. Additionally, we did not test the confounding effect of sites, due to their large number. In general, country and site are not accounted for in studies applying GMM. Another limitation by the study design is that all patients in the study received open label sertraline or paroxetine and a double-blinded placebo. It is possible that addition of a placebo medication resulted in a larger treatment response compared to a treatment with sertraline or paroxetine alone, due to possible expectations that the placebo medication may have been a new active drug. The aim of this study was not to investigate the general efficacy of SSRI treatment but to examine the response. However, we find it worth noticing that the response rate of 58.5 % in this study is within the range of SSRI response rates observed in studies that significantly differentiated from placebo. In addition, it would be interesting to investigate treatment responses to other types of antidepressants, such as the serotonin-norepinephrine reuptake inhibitors (SNRI) venlafaxine extended-release, which has shown to reduce PTSD symptoms compared to placebo (Davidson et al., 2006).

It was beyond the scope of this paper to investigate if the predictors affected class-specific growth parameters. Including class-specific growth parameters could have been used to better understand the predictors’ effect within class and Muthén (2003, 2004) argues that it may also result in a more reliable model. However, it would also result in a more complex model with a smaller ratio between model parameters and sample size.

Future research studies with larger sample sizes could attempt to replicate our findings and could also fully explore the spectrum of trajectories. It is possible that the 4-class model is the best model in a larger sample. It would also be interesting to conduct exploratory studies with a wider range of predictors. In this context, elevated startle response has shown potential (Arikan et al., 2006) but also genetic markers, such as genes related to drug metabolism or polygenic risk scores for the related
comorbidities would be of high interest. In our opinion, GMM has been a good method to investigate the heterogeneity of treatment response. However, we also recommend using GCM to investigate the overall population level response. A better understanding of the population level response can help us better understand and interpret the findings from the GMM.

**Formatting of funding sources**

The data used in this research is from a study conducted in a collaboration between Otsuka Pharmaceutical Development & Commercialization Inc. and H. Lundbeck A/S. Anne Krogh Nohr is funded by Innovation Fund Denmark (grant no. 8053-00004A). Anders Albrechtsen is supported by the Lundbeck Foundation (R215-2015-4174) and Novo Nordisk Foundation (NNF20OC0061343).

**Limitations**

The analyzed data was from the prospective phase of a clinical trial. The patients were treated with open-label sertraline or paroxetine and a double-blinded placebo.

**Contributions**

Anne Krogh Nohr was responsible for the data analysis and wrote the first draft of the manuscript. Hans Eriksson, Raimund Buller, and Mary Hobart designed the study and wrote the protocol. Ida Moldtke, Anders Albrechtsen, and Stinus Lindgreen supervised the statistical analysis. All authors contributed to and have approved the final manuscript.

**Declaration of Competing Interest**

Anne Krogh Nohr is employed at H. Lundbeck A/S and Hans Eriksson, Raimund Buller, and Stinus Lindgreen were employees when this study was conducted. Mary Hobart is an employee of Otsuka. All other authors declare no conflicts of interest.

**Acknowledgement**

We would like to thank Christoph Von der Goltz, who read and gave his perspective on this article.

**Supplementary materials**

Supplementary material associated with this article can be found in the online version, at doi:10.1016/j.psychres.2021.113964.

**References**


Albrechtsen is supported by the Lundbeck Foundation (R215-2015-4174) and Novo Nordisk Foundation (NNF20OC0061343).


