Genes in treatment: Polygenic risk scores for different psychopathologies, neuroticism, educational attainment and IQ and the outcome of two different exposure-based fear treatments

André Wannemüller, Robert Kumsta, Hans-Peter Jöhren, Thalia C. Eley, Tobias Teismann, Dirk Moser, Christopher Rayner, Gerome Breen, Jonathan Coleman, Svenja Schaumburg, Simon E. Blackwell & Jürgen Margraf


To link to this article: https://doi.org/10.1080/15622975.2021.1907708
ABSTRACT

Objectives: Evidence for a genetic influence on psychological treatment outcome so far has been inconsistent, likely due to the focus on candidate genes and the heterogeneity of the disorders treated. Using polygenic risk scores (PRS) in homogenous patient samples may increase the chance of detecting genetic influences.

Methods: A sample of 342 phobic patients treated either for clinically relevant dental fear (n = 189) or other (mixed) phobic fears (n = 153) underwent highly standardised exposure-based CBT. A brief five-session format was used to treat dental fear, whereas longer multi-session treatments were used with the mixed-fear cohort. PRS were calculated based on large genetic studies of Neuroticism, Educational Attainment (EA), Intelligence, and four psychopathology domains. We compared PRS of post-treatment and follow-up remitters and non-remitters and regressed PRS on fear reduction percentages.

Results: In the dental fear cohort, EA PRS were associated with treatment outcomes, i.e. dropout, short- and long-term remission state, fear reduction, and attendance of subsequent dental appointments. In the mixed fear treatment cohort, no gene effects were observable.

Conclusions: Results indicate the importance of EA-related traits for outcomes following brief, but not long, standardised exposure-based CBT. Such use of PRS may help inform selection and tailoring of treatments.

ARTICLE HISTORY

Received 24 August 2020
Revised 14 January 2021
Accepted 13 February 2021

KEYWORDS

Therapygenetics; exposure treatment; genetics; polygenic scores; phobias

Introduction

Among cognitive behavioural treatments (CBT), exposure-based strategies are considered the gold standard for treating fear and avoidance behaviour. Exposure therapy has been shown to be highly effective, whether applied alone or in combination with cognitive restructuring or somatic coping strategies (Norton and Price 2007; Hofmann and Smits 2008). However, there is a substantial inter-individual variation in treatment response and long-term treatment efficacy. While many patients show a complete recovery, others continue to experience significant levels of symptoms after treatment (Loerinc et al. 2015) or display a (partial) return of their fear symptoms over time (Craske and Mystkowski 2006). Therefore, in order to further improve the effectiveness of psychotherapeutic interventions, it is necessary to identify psychological and biological factors that influence their outcomes.

The research field termed ‘therapygenetics’ is still relatively new and aims to identify genetic markers that contribute to the observed variance in treatment outcomes. In the future, within a ‘stratified medicine’ (Trusheim et al. 2007) framework, such markers could be used to match patients to the treatment formats from which they are most likely to benefit. So far, most therapygenetic studies have focussed on candidate markers in a limited number of genes, with inconclusive results in terms of their predictive value as well as the direction of effect (see Lester and Eley 2013 for the latest review). This is not surprising, given
the complex, polygenic nature of disorders like anxiety or depression (Smoller 2016) for which single genetic variants explain extremely small portions of variance (Gratten et al. 2014). The same will also likely be true for genetic variants predicting therapy outcomes. Existing studies are thus likely to be underpowered to detect single genetic variants or may not have looked for genetic markers of treatment success at the right gene loci (see Duncan and Keller 2011 for a summary of problems associated with the candidate gene approach in psychiatry).

In addition to the narrow focus on candidate genes, therapygenetic studies also face several other difficulties. In order to increase sample sizes, most studies combine data from different sources. This often means that samples are highly heterogeneous in terms of treatment type, content, dose, and form of delivery (Lester and Eley 2013). One approach to improve statistical power to detect gene by treatment effects is to use more homogeneous datasets, i.e. data from highly standardised treatments, applied among patients with a specific diagnosis (Beevers and McGeary 2012), in a standardised treatment setting (see Wannemüller et al. 2018) for a recent candidate gene approach in a highly standardised sample.

A possible way to circumvent some of the described problems of candidate gene approaches may consist in applying so-called polygenic risk scores (PRS). PRS represent the multiplied count of all genetic variants with phenotype-specific predetermined effect sizes and reflect the genetic propensity of an individual towards this phenotype through a single summary score (see Wray et al. 2014 for a review). The summary statistics required for the calculation of independent PRS result from previous genome-wide association studies (GWAS) that normally include tens of thousands of case and control participants and in which allele frequencies for each common genetic variant are tested for their association with the respective trait of interest. This allows for the calculation of an effect size and a p-value for each tested variant.

Today, GWAS exist for many mental disorders, making it possible to calculate PRS for them. It can be expected that the usefulness of PRS as screening instruments to estimate the disorder risk will increase with the size of the available GWAS in the future. Since the risk of developing a mental disorder is only partly due to genetic factors and PRS can only capture parts of the absolute genetic risk, their predictive utility will always be limited (see Anderson et al. 2019; Murray et al. 2021 for an overview concerning limitations of using polygenic risk scoring in the field of mental health). However, PRS for mental disorders could represent useful decision-making aids concerning the indication and allocation of patients to ‘individually tailored treatments’, if they were to prove predictive of the outcome of (different) treatments (see Murray et al. 2021 for a review). Initial psychopharmacological findings point in that direction. For example, it has been shown that the response to clozapine treatment in patients with schizophrenia was associated with the level of their schizophrenia PRS (Ruderfer et al. 2016; Zhang et al. 2019). However, it has not yet been investigated whether PRS for the treated primary disorder are also associated with the outcome of psychological treatment and to what extent this also applies to the genetic disposition for comorbid disorders. The latter can be expected because co-morbidities have been evidenced to reduce the outcome expectation of psychological treatments (e.g. Steketee et al. 2001; Wolitzky-Taylor et al. 2012; Halldorsdottir and Ollendick 2014). In the long-term, therapygenetic research will benefit from well-powered GWAS of ‘treatment response’. However, so far only one rather small GWAS exists for this phenotype (Coleman et al. 2016) and no large GWAS consortia have been constituted on this topic yet. This might be due to several difficulties concerning a precise definition of this phenotype and the necessity of long-term follow-up observations in the patients (see Murray et al. 2021 for details). As long as there are no large GWAS results, it might be useful to use polygenic risk scores in order to examine genetic overlap between psychological therapy outcomes and related traits as proposed by Uher (2008).

In this study, we applied PRS for four different mental disorders [General Anxiety Disorder (GAD), Major Depressive Disorder (MDD), Schizophrenia (SCZ) and Autism Spectrum Disorder (ASD)] as predictors of baseline symptom severity as well as short- and long-term treatment outcomes in two cohorts of phobic patients. Patients were treated either for clinically relevant dental fear with a highly standardised brief five-session exposure-based fear treatment \( n = 189 \) or for other phobic disorders (e.g. specific phobias or agoraphobia) with standardised multi-session phobia-specific exposure-based CBT \( n = 153 \). Additionally, we included the polygenic score for Neuroticism – a personality trait known to be associated with poor outcomes following psychological treatment and a tendency to relapse (Tyrer et al. 1992; Schuurmans et al. 2009).

In addition to these PRS hypothesised to limit the response to treatment we included the PRS for
Educational attainment (EA). EA has been shown to associate with psychological traits (Erickson et al. 2016), treatment-relevant behaviours such as premature drop-out (Wierzbicki and Pekarik 1993; Roseborough et al. 2016) and the outcome of psychological treatment (Ehlers et al. 2005). However, in contrast to comorbidity the association between EA and outcome parameters was positive. Therefore, we examined associations between treatment outcomes and the EA polygenic score, as well as an IQ polygenic score. We included the IQ polygenic score to allow for comparison with the education polygenic score, in order to investigate whether genetic influences on treatment outcomes were more strongly influenced by cognitive abilities alone, or whether education-related traits might be more relevant. The polygenic score for body mass index was included in our analyses as a control score as it does not directly refer to a psychological trait. We hypothesised that in both cohorts polygenic scores for mental disorders and Neuroticism would be associated with both baseline symptom severity and poorer treatment outcomes whereas EA PRS should rather be associated with favourable treatment outcomes.

Materials and methods

Participants

Our sample consisted of 342 phobic patients (66.7% female) with a mean age of 37.8 ± 11.2 years. Patients belonged to one of two treatment cohorts that differed in relation to the content of the phobic fears and the treatment duration.

Dental fear cohort

The dental fear cohort consisted of 189 participants with pathological dental fear and had predominantly middle-European ancestry (>96% of the sample). Primary diagnoses based on DSM-IV (APA 2000) included dental phobia (DP; n = 159; 84.1%), blood-injury-injection phobia (n = 13; 6.9%), panic disorder with or without agoraphobia (n = 6; 3.2%), social phobia (n = 4; 2.1%) and posttraumatic stress disorder (n = 3; 1.6%). One patient was diagnosed with hypochondridia (0.5%), another with specific phobia regarding contamination (0.5%) and another with agoraphobia without panic disorder (0.5%). The primary diagnosis was unclear in one patient (0.5%) as he showed mixed dental phobic and blood-injury-injection phobic symptoms. Besides primary diagnoses, 78 patients (41.3%) had at least one comorbidity. The most common comorbidities were specific phobias (n = 38), substance abuse/dependency (excluding nicotine, n = 14) and major depression (n = 12). All patients were recruited via a dental clinic, which specialises in treating individuals with dental fear. As part of the routine intake assessment, dental fear is screened using a German questionnaire, the ‘Hierarchischer Angstfragebogen’ (HAF; Jöhren 1999; see below). Patients exceeding a score of 38 on the HAF were assigned to a Treatment Centre for Dental Fear (TCDF) a psychological unit specialising in diagnosing and treating dental fear. In the TCDF licenced psychotherapists experienced in treating pathological dental fear conducted the clinical diagnostics and subsequent treatment. Patients who started their psychological treatment between April 2012 and June 2014 were included in the study.

Mixed fear cohort

Patients of the mixed fear cohort were N = 153 individuals with predominantly middle-European ancestry (>92%). Participants’ primary diagnosis was a specific phobia in 39.2% of the cases (n = 60) with the ‘animal’ (n = 17; 11.1% of total) and ‘environmental’ (n = 18; 11.8%) subtypes being the most frequent subtypes, followed by the ‘situational’ (n = 12; 7.8%), ‘other’ (n = 10; 6.5%) and ‘blood-injection-injury’ subtypes (n = 3; 2.0%). In all other participants (n = 93; 60.8%), the primary diagnosis was agoraphobia. In the vast majority of cases (n = 87; 52.2% of total) agoraphobia was associated with a panic disorder. Only in 7 patients (4.6% of total), was agoraphobia present without a history of panic disorder. In addition to their primary diagnosis, 54 patients (35.3%) had at least one comorbidity, with specific phobias (n = 31) and social phobia (n = 12) most common. All patients aged between 18 and 70 years who sought treatment between December 2011 and November 2014 in an outpatient university clinic, for either agoraphobia or specific phobia were invited to participate in this study. Participants were excluded if they had comorbid bipolar disorder, psychotic disorder, alcohol/substance abuse or dependency (within the past 3 months, excluding nicotine). Prominent risk of self-harm, organic mental disorder and concurrent psychotherapeutic or psychopharmacological treatment also led to exclusion.

All patients were informed that the clinic regularly conducts research and provided informed consent prior to participation, and provided specific consent for genome-wide genotyping. No compensation was given to clients for participation. The study as
described here was reviewed and approved by the local Ethics Committee (number 011/2012).

**Treatments**

**Five-session treatment**

Treatment for the dental fear cohort treatment was delivered via five weekly individual sessions according to a highly standardised treatment manual (Sartory and Wannemüller 2010; Wannemüller et al. 2015). In the first session, a semi-structured clinical interview (Mini-DIPS; Margraf 1994) was conducted to confirm the diagnosis underlying their pathological dental fear as well as existing comorbidities, and psychoeducational information was provided. In the second session, patients were introduced to bodily and cognitive coping strategies (applied relaxation, breathing techniques, helpful thoughts) to be practiced at home and applied to counter bodily fear symptoms in dental situations. The third to fifth sessions all consisted of exposure exercises including sound, video, and in-vivo exposure. Following the last treatment session, the therapist ascertained whether the patient was remitted, i.e. no longer meeting diagnostic criteria for the primary diagnosis underlying their dental fear. Nine months ($M = 9.54 \pm 5.06$) after treatment completion, patients were invited for a follow-up (FU) assessment containing the same questionnaire measures as presented at the pre- and post-treatment assessments. Moreover, some additional treatment-relevant information was collected. Patients were asked how often they had attended a dental appointment in the post-treatment follow-up interval and to rate the intensity and frequency with which they practiced and applied the learned coping strategies during dental treatments. Furthermore, they rated the helpfulness of each of these strategies. A postgraduate psychologist who was blind to the post-treatment diagnostic status of the respective patient conducted a clinical interview using the Mini-DIPS and ascertained whether the patient was remitted or non-remitted. Ninety-three patients (64.6% of treatment completers) were available for FU-assessment.

**Multi-session treatment**

Individual cognitive-behavioural treatment for the mixed fear cohort was conducted by postgraduate clinical psychologists with regular supervision (including use of audio-visual recordings) by experienced senior clinicians. Participants with a primary diagnosis of agoraphobia were randomly assigned to one of two treatment formats, a cognitive-behavioural treatment or to an exposure-alone condition without any element of cognitive restructuring. All participants received five preliminary sessions covering diagnostics based on a semi-structured clinical interview (DIPS; Schneider and Margraf 2006) and psychoeducation before starting therapy. Participants with a specific phobia were treated in a long-term program of up to twenty-five individual sessions of in vivo exposure. Six months ($M = 5.70 \pm 1.37$) after completing the treatment, patients were invited for a follow-up (FU) assessment containing the same questionnaire measures as presented at the pre- and post-treatment assessments. A postgraduate psychologist who was blind to the post-treatment diagnostic status of the respective patient assessed their diagnostic status based on the DIPS. One hundred patients (65.4% of the original sample and 90.1% of treatment completers) were available for FU-assessment.

**Outcome measures**

**Dental fear cohort**

In order to assess the anticipatory and subjective components of dental fear the Dental Anxiety Scale (DAS; Corah 1969) and the HAF (Jöhren 1999) were used. The DAS describes four and the HAF eleven dental-related situations, and patients rate how anxious they would feel in these situations on a 5-point Likert-scale ($1 = \text{relaxed}; 5 = \text{extremely anxious}$). Internal consistency indices were (Cronbach’s $\alpha$) $\alpha = .84$ (DAS) and $\alpha = .91$ (HAF) in the present sample. Dysfunctional dental-related cognitions were assessed with the Dental Cognition Questionnaire (DCQ; De Jongh et al. 1995). The DCQ consists of 38 dichotomous items (yes/no) measuring dysfunctional dental-related cognitions that might emerge prior to (Items 1–14) or during (Items 15–38) dental surgery. Internal consistency was good in the present sample, $\alpha = .87$. Experience of control in dental situations was assessed using the Iowa Dental Control Index Revised (IDCI-R; Brunsman et al. 2003). This consists of nine items measuring ‘felt control’ (four items) and ‘desired control’ (five items) during dental surgery. Items were answered on a 5-point Likert-scale ($1 = \text{no control at all}; 5 = \text{total control}$). In the present sample Cronbach’s $\alpha$ was $\alpha = .82$ for the ‘desired control’ and $\alpha = .75$ for the felt control scale.

In order to construct a robust indicator of dental fear including all fear facets rather than looking at one primary outcome we created a composite score consisting of the mean percentage of maximum score of all dental-fear related questionnaires (DAS, HAF, DCQ,
When a patient's primary diagnosis was agoraphobia (with or without a history of panic) the Mobility Inventory (MI; Chambless et al. 1985) was used to assess agoraphobic symptoms. The MI is a self-report questionnaire that measures the degree to which 27 situations are avoided. Items are scored from 1 (never avoid the situation) to 5 (always avoid the situation), with the mean of all items used as the total score. For this study, only the ratings for the "alone" subscale were used. The internal consistency of the MI in this sample was $\alpha = .94$. For patients with agoraphobia, we used the mean percentage score of the MI (alone scale) at pre-treatment for analyses of agoraphobic fear at baseline (pre-treatment fear level). For analyses of outcomes following treatment, we calculated the mean percentage change on the MI (alone scale) at post-treatment and at 9-month follow-up. In order to assess specific fear symptoms in the phobic participants, we used two separate 9-point rating-scales measuring subjective fear and avoidance (0 = no fear/no avoidance; 8 = very intense fear/total avoidance).

First, we asked participants to state which situational or object-related fear (e.g. fear of snakes, dogs, spiders, fear of heights, fear of flying, fear of blood and injections etc.) they were seeking treatment for. They were then asked to rate their fear and avoidance levels using the respective DIPS scales. For the analyses of phobic fear at baseline in individuals with a specific phobia, we calculated the mean percentage of a maximum score of the averaged fear and avoidance scale (pre-treatment fear level). For outcome analyses following treatment, we calculated the mean percentage change averaged across both scales at post-treatment and 6-months follow-up. Again, in the case of 'intention-to-treat' data, post-treatment change was calculated using LOCF as a data imputation method in individuals who dropped out prematurely (i.e. 20.6% of the dental fear cohort, see Table 1). Patients who completed the treatment but did not provide post-treatment questionnaire data ($n = 28$, 14.8%) were excluded from the regression analyses. Patients who terminated the treatment consensually ($n = 6$, 3.2%) were excluded from the regression analyses and group comparisons since for these patients neither questionnaire data nor remission state were available.

**Mixed fear cohort**

When a patient's primary diagnosis was agoraphobia (with or without a history of panic) the Mobility Inventory (MI; Chambless et al. 1985) was used to assess agoraphobic symptoms. The MI is a self-report questionnaire that measures the degree to which 27 situations are avoided. Items are scored from 1 (never avoid the situation) to 5 (always avoid the situation), with the mean of all items used as the total score. For this study, only the ratings for the "alone" subscale were used. The internal consistency of the MI in this sample was $\alpha = .94$. For patients with agoraphobia, we used the mean percentage score of the MI (alone scale) at pre-treatment for analyses of agoraphobic fear at baseline (pre-treatment fear level). For analyses of outcomes following treatment, we calculated the mean percentage change on the MI (alone scale) at post-treatment and 9-month follow-up. In order to assess specific fear symptoms in the phobic participants, we used two separate 9-point rating-scales measuring subjective fear and avoidance (0 = no fear/no avoidance; 8 = very intense fear/total avoidance).

First, we asked participants to state which situational or object-related fear (e.g. fear of snakes, dogs, spiders, fear of heights, fear of flying, fear of blood and injections etc.) they were seeking treatment for. They were then asked to rate their fear and avoidance levels using the respective DIPS scales. For the analyses of phobic fear at baseline in individuals with a specific phobia, we calculated the mean percentage of a maximum score of the averaged fear and avoidance scale (pre-treatment fear level). For outcome analyses following treatment, we calculated the mean percentage change averaged across both scales at post-treatment and 6-months follow-up. Again, in the case of 'intention-to-treat' data, post-treatment change was calculated using LOCF as a data imputation method in individuals who dropped out prematurely ($n = 27$, i.e. 17.6% of the mixed fear cohort, see Table 1). Patients who completed the treatment but did not provide post-treatment questionnaire data ($n = 15$, 9.9%) where excluded from regression analyses. In patients who terminated the treatment consensually ($n = 15$, 9.9%) the Mobility Inventory (MI; Chambless et al. 1985) was used to assess agoraphobic symptoms. The MI is a self-report questionnaire that measures the degree to which 27 situations are avoided. Items are scored from 1 (never avoid the situation) to 5 (always avoid the situation), with the mean of all items used as the total score. For this study, only the ratings for the "alone" subscale were used. The internal consistency of the MI in this sample was $\alpha = .94$. For patients with agoraphobia, we used the mean percentage score of the MI (alone scale) at pre-treatment for analyses of agoraphobic fear at baseline (pre-treatment fear level). For analyses of outcomes following treatment, we calculated the mean percentage change on the MI (alone scale) at post-treatment and 9-month follow-up. In order to assess specific fear symptoms in the phobic participants, we used two separate 9-point rating-scales measuring subjective fear and avoidance (0 = no fear/no avoidance; 8 = very intense fear/total avoidance).

First, we asked participants to state which situational or object-related fear (e.g. fear of snakes, dogs, spiders, fear of heights, fear of flying, fear of blood and injections etc.) they were seeking treatment for. They were then asked to rate their fear and avoidance levels using the respective DIPS scales. For the analyses of phobic fear at baseline in individuals with a specific phobia, we calculated the mean percentage of a maximum score of the averaged fear and avoidance scale (pre-treatment fear level). For outcome analyses following treatment, we calculated the mean percentage change averaged across both scales at post-treatment and 6-months follow-up. Again, in the case of 'intention-to-treat' data, post-treatment change was calculated using LOCF as a data imputation method in individuals who dropped out prematurely ($n = 27$, i.e. 17.6% of the mixed fear cohort, see Table 1). Patients who completed the treatment but did not provide post-treatment questionnaire data ($n = 15$, 9.9%) where excluded from regression analyses. In patients who terminated the treatment consensually ($n = 15$, 9.9%) the Mobility Inventory (MI; Chambless et al. 1985) was used to assess agoraphobic symptoms. The MI is a self-report questionnaire that measures the degree to which 27 situations are avoided. Items are scored from 1 (never avoid the situation) to 5 (always avoid the situation), with the mean of all items used as the total score. For this study, only the ratings for the "alone" subscale were used. The internal consistency of the MI in this sample was $\alpha = .94$. For patients with agoraphobia, we used the mean percentage score of the MI (alone scale) at pre-treatment for analyses of agoraphobic fear at baseline (pre-treatment fear level). For analyses of outcomes following treatment, we calculated the mean percentage change on the MI (alone scale) at post-treatment and 9-month follow-up. In order to assess specific fear symptoms in the phobic participants, we used two separate 9-point rating-scales measuring subjective fear and avoidance (0 = no fear/no avoidance; 8 = very intense fear/total avoidance).

First, we asked participants to state which situational or object-related fear (e.g. fear of snakes, dogs, spiders, fear of heights, fear of flying, fear of blood and injections etc.) they were seeking treatment for. They were then asked to rate their fear and avoidance levels using the respective DIPS scales. For the analyses of phobic fear at baseline in individuals with a specific phobia, we calculated the mean percentage of a maximum score of the averaged fear and avoidance scale (pre-treatment fear level). For outcome analyses following treatment, we calculated the mean percentage change averaged across both scales at post-treatment and 6-months follow-up. Again, in the case of 'intention-to-treat' data, post-treatment change was calculated using LOCF as a data imputation method in individuals who dropped out prematurely ($n = 27$, i.e. 17.6% of the mixed fear cohort, see Table 1). Patients who completed the treatment but did not provide post-treatment questionnaire data ($n = 15$, 9.9%) where excluded from regression analyses. In patients who terminated the treatment consensually ($n = 15$, 9.9%) the Mobility Inventory (MI; Chambless et al. 1985) was used to assess agoraphobic symptoms. The MI is a self-report questionnaire that measures the degree to which 27 situations are avoided. Items are scored from 1 (never avoid the situation) to 5 (always avoid the situation), with the mean of all items used as the total score. For this study, only the ratings for the "alone" subscale were used. The internal consistency of the MI in this sample was $\alpha = .94$. For patients with agoraphobia, we used the mean percentage score of the MI (alone scale) at pre-treatment for analyses of agoraphobic fear at baseline (pre-treatment fear level). For analyses of outcomes following treatment, we calculated the mean percentage change on the MI (alone scale) at post-treatment and 9-month follow-up. In order to assess specific fear symptoms in the phobic participants, we used two separate 9-point rating-scales measuring subjective fear and avoidance (0 = no fear/no avoidance; 8 = very intense fear/total avoidance).

First, we asked participants to state which situational or object-related fear (e.g. fear of snakes, dogs, spiders, fear of heights, fear of flying, fear of blood and injections etc.) they were seeking treatment for. They were then asked to rate their fear and avoidance levels using the respective DIPS scales. For the analyses of phobic fear at baseline in individuals with a specific phobia, we calculated the mean percentage of a maximum score of the averaged fear and avoidance scale (pre-treatment fear level). For outcome analyses following treatment, we calculated the mean percentage change averaged across both scales at post-treatment and 6-months follow-up. Again, in the case of 'intention-to-treat' data, post-treatment change was calculated using LOCF as a data imputation method in individuals who dropped out prematurely ($n = 27$, i.e. 17.6% of the mixed fear cohort, see Table 1). Patients who completed the treatment but did not provide post-treatment questionnaire data ($n = 15$, 9.9%) where excluded from regression analyses. In patients who terminated the treatment consensually ($n = 15$, 9.9%)
9.9%) questionnaire data was available in this cohort and these patients were only excluded from group comparisons concerning remission.

**Mental distress (depression, anxiety, stress)**

In order to assess general symptoms of psychopathology we used the short form of the Depression-Anxiety-Stress Scale (DASS-21; Lovibond and Lovibond 1995) in both cohorts. This self-rating questionnaire consists of three seven-item scales, assessing the levels of depression, anxiety and stress. Items are answered on a four-point scale (0 = not at all applicable; 3 = highly applicable). Very good psychometric properties of the DASS-21 have been reported (Antony et al. 1998). Here, we found an internal consistency of the total score of Cronbach’s $\alpha = .93$ (depression: $\alpha = .90$; stress: $\alpha = .85$; anxiety $\alpha = .82$).

**Genotyping procedures**

DNA from most participants ($n = 364$) was extracted from blood, with the rest ($n = 68$) obtained from saliva samples by routine desalting methods. Saliva samples were taken from patients who refused to have blood drawn due to high levels of blood-injury-injection fear. In order to avoid that blood drawing would have an influence on fear ratings, completion of questionnaires and blood drawings took place on different days. Genotyping of the participants was performed using the Illumina PsychChip microarray (Illumina, USA), see (Rayner et al. 2019) for a detailed description. In brief, variants were excluded if they were missing in >1% of participants, if they had a minor allele frequency < .05, or if they deviated substantially from Hardy–Weinberg equilibrium ($p < 10^{-5}$). Individuals were included only if they had genotype calls for <99% of variants, had concordant phenotypic and genotypic sex (X chromosome heterozygosity $F$ statistic: males > .8 and females < .2), were unrelated to other individuals (identity by descent: IBD > .1875) and showed no evidence of contamination (pairwise IBD and genome-wide heterozygosity $F$ statistic < 3 SD from the cohort means).

**Polygenic scoring**

Polygenic scores were computed using publicly available GWAS summary statistics for the respective trait of interest, i.e. Schizophrenia (SWGPGC 2014), Major Depressive Disorder (MDD; Wray et al. 2018), Anxiety Disorders (GAD; Purves et al. 2017), Autism Spectrum Disorders (ASD; The ASDWPGPC 2017), Neuroticism (NEUR; Nagel et al. 2018), Intelligence Quotient (IQ; Savage et al. 2018), Educational Attainment (EA; Okbay et al. 2016) and BMI as a control dimension (UK Biobank 2018).

Polygenic scores for each trait were calculated based on seven $p$-value thresholds that mark the significance level by which a SNP has to be associated with the trait of interest in order to be included in the PRS-calculation ($p$-value $< 10^{-8}$; $< .0001$; $< .05$; $< .1$; $< .2$; $< .5$; $< 1.0$). For further analyses, we selected the $p$-value thresholds based on their associations with trait indicators within our total sample. In case of Neuroticism (DASS-Depression and Anxiety), GAD (DASS-Anxiety) and MDD (DASS-Depression), PRS thresholds were selected based on their correlations to the respective DASS scales. For the EA and IQ PRS, the threshold was selected based on maximal correlation to academic attainments and for the BMI we observed correlation levels between PRS and the phenotypical BMI (for an overview of all PRS-trait correlations and selected $p$-value thresholds see Supplementary Table 1). For Schizophrenia and ASD there was no phenotypical trait indicator available in our sample. We therefore selected a $p$-value threshold of $p < .2$ as in the Schizophrenia GWAS (SWGPGC 2014) risk alleles at a $p < .02$ threshold explained the largest proportion of phenotypical variance.

**Statistical analyses**

We compared polygenic scores of treatment remitters and non-remitters for each domain using multivariate ANOVAs with age, sex, ten principal components derived from the genotype data (to control for population stratification), and cohort (in case of pooled data analyses) as covariates to determine whether polygenic scores differed between remitters and non-remitters.

Furthermore, we used linear hierarchical linear regression analyses in order to predict pre-treatment fear levels, the percentage of pre to post, and pre to follow-up fear reduction respectively, and the change in subjective fear percentage in the post to follow-up interval from genotype data. Block 1 included the covariates age, sex, the ten principal components (to control for population stratification), and cohort (in case of pooled data analyses). Block 2 always contained the PRS at the selected $p$-value thresholds. Our final sample size of $N = 295$ in the regression analyses containing 8 different PRS and 13 control variables (age, sex, cohort and 10 principal components) would provide us with 80% power to detect effect sizes ($R^2$) of approximately .07 (small-to-medium) in our main
analyses concerning whole model fit. According to Green (1991), the required sample size that would provide us with an 80% power to find medium effect sized beta-coefficients ($R^2=0.07$) under these conditions is $N=125$.

All analyses were conducted using the IBM Statistics SPSS 24 software package.

Results

Sample description and treatment outcome

In terms of sociodemographic variables, the patients in the two cohorts differed only in their educational level, $t=4.37$, $p<.001$, with more years of academic education in the mixed fear cohort. Moreover, pre-treatment fear levels differed between the cohorts, $t=7.28$, $p<.001$, with higher levels in the dental fear cohort. Treatment length was of course shorter in the dental fear cohort, $t=20.00$, $p<.001$, and resulted in comparably less fear reduction at both post-treatment, $t=5.98$, $p<.001$, and follow-up assessment, $t=3.63$, $p<.001$, see Table 1.

Associations between polygenic scores and pre-treatment fear levels

Results of hierarchical linear regression controlling for age, sex, population stratification, and cohort showed that none of the polygenic scores examined were significantly associated with pre-treatment fear levels in the whole sample. The same was true within both individual fear cohorts, see Supplementary Table II.

A regression analysis using pre-treatment mental distress levels (DASS-21 total score) as dependent variable yielded the PRS for neuroticism $b^*=.142$, $p=.022$ and educational attainment $b^*=-.151$, $p=.026$ as strongest predictors in the whole sample. However, the predictive effects differed between the cohorts. In the dental fear cohort only EA PRS predicted distress, $b^*=-.241$, $p=.019$, whereas in the mixed fear cohort neuroticism PRS was the strongest predictor, $b^*=.197$, $p=.029$, see Supplementary Table III.

Associations between PRS, post-treatment remission status and fear reduction

A MANOVA containing the polygenic scores at the selected $p$-value thresholds for each trait demonstrates significantly higher EA polygenic scores in post-treatment remitters compared to non-remitters, $F=7.69$, $p=.006$, $\eta^2=.03$ in the whole sample, see Figure 1. Separate post hoc MANOVAs for each cohort showed a large difference in relation to EA-PRS between remitters and non-remitters in the dental fear cohort treated in five-sessions, $F=13.23$, $p=.0004$, $\eta^2=.07$ as well as - to a much smaller extent - GAD PRS, $F=3.98$, $p=.048$, $\eta^2=.02$, with no significant differences found for the mixed fear cohort, see Figure 1.

A post hoc analysis in the dental fear cohort comparing PRS between post-treatment remitters ($n=102$), patients who completed the treatment but
Figure 1. Top. Estimated mean polygenic scores of patients who were remitted at post treatment and those who were not remitted or dropped out prematurely (controlling for age, sex, population stratification and cohort). Centre. Estimated mean polygenic scores of patients who were remitted at post-treatment and those who were not remitted or dropped out prematurely in the five-session treatment (controlling for age, sex and population stratification). Bottom. Estimated mean polygenic scores of patients who were remitted at post-treatment and those who were not remitted or dropped out prematurely in the multi-session treatment (controlling for age, sex and population stratification). EA: Educational Attainment; IQ: Intelligence Quotient; NEUR: Neuroticism; SCHIZ: Schizophrenia; GAD: General Anxiety Disorder; MDD: Major Depression Disorder; BMI: Body Mass Index; error bars indicate standard errors of mean; * contains also patients who were not remitted at post-treatment assessment and did not attend follow-up assessment.
did not remit (n = 39) and those who dropped out prematurely (n = 37) demonstrated that EA PRS in remitters was higher $F = 7.77, p = .001, \eta^2 = .09$ compared to both drop-outs ($p = .004$) and non-remitters ($p = .001$).

Consistently, a regression analysis predicting pre to post-treatment percentage fear reduction (intention-to-treat) from polygenic scores revealed the EA PRS to significantly predict fear reduction in the dental fear cohort, $b^* = .270, p = .008$, see Table 2. Moreover, higher neuroticism PRS predicted less post-treatment fear reduction, $b^* = -.213, p = .026$, albeit to a much lesser extent.

Splitting patients in the dental fear cohort into those individuals with comparably high education levels (A-level or higher, $n = 60$) and those with relatively below-average education levels (less than A-level, $n = 120$) demonstrated that in the latter group EA-PRS explained 10% of the existing variance between remitters and non-remitters, $F = 11.73, p = .001, \eta^2 = .10$ whereas no significant proportion of variance was explained in the higher educated group, $F = .30, p = .58, \eta^2 < .01$, see Supplementary Figure I. Moreover, only in the lower educated group EA-PRS were predictive of post-treatment fear reduction ($b^* = .268, p = .016$) whereas in the higher educated group this was not the case ($b^* = .201, p = .194$).

**Associations between PRS, follow-up remission status and fear reduction**

At follow-up a similar pattern emerged as observed for post-treatment assessment. A MANOVA comparing remitters and non-remitters (including patients who were not remitted at post treatment and did not attend FU-assessment) again showed a significant difference between remitters and non-remitters in terms of EA polygenic score in the dental fear cohort, $F = 8.76, p = .004$, see Figure 2.

Regression analyses again yielded EA PRS as the sole predictor of follow-up fear reduction in the dental fear cohort, $b^* = .234, p = .016$, see Table 2.

To further validate the results in the dental fear cohort we correlated the EA PRS with objective behaviours as indicators for therapy success, i.e. how often participants attended dental treatments within the post-treatment to follow-up interval. Again, all analyses controlled for age, sex and population stratification. PRS for EA was positively associated with the number of dental appointments attended in the post-treatment to follow-up interval, $r = .295, p = .011$, see Supplementary Figure II.

**Associations between PRS and changes in fear in the post-treatment to follow-up interval**

In order to test these associations, we regressed PRS on the difference score between fear reduction at post-treatment and follow-up ($\Delta \%$) in follow-up returners, see Supplementary Table IV. None of the PRS scores based on the selected $p$-value thresholds significantly predicted change in fear in the whole sample or in either individual subcohort.

**Discussion**

This study aimed to investigate whether polygenic scores for four psychopathology domains (GAD, MDD, ASD, Schizophrenia), as well as those for Neuroticism, IQ and Educational Attainment influenced the outcomes of either a highly standardised brief exposure-based treatment of dental fear or longer-term exposure-based treatment of other phobias. Furthermore, we aimed to identify polygenic scores that are associated with baseline severity of phobic fear and symptom distress.

In contrast to our expectations none of the analysed polygenic dispositions – regardless of whether psychopathology-related or not – yielded a direct influence on pre-treatment phobic fear levels in our patients. However, PRSs were associated with the levels of pre-treatment distress (DASS-21) in the participants. Interestingly, the PRS that were most strongly associated with distress levels differed between cohorts. While in the mixed fear cohort the PRS for neuroticism, i.e. a disposition related to pleiotropic psychopathology (Nagel et al. 2018), was most strongly related to distress levels, in the dental fear cohort only the PRS for educational attainment was a significant predictor of pre-treatment distress. It is possible that differences between the cohorts, to be discussed later, may have contributed to this discrepancy.

For participants in the dental fear cohort, who were treated in five highly standardised exposure-based treatment sessions, we could trace the previously reported associations between educational attainment and outcome of psychological treatments (Ehlers et al. 2005) down to a genetic level. In this cohort, we observed associations between the polygenic disposition for EA and premature drop-out, remission status, and phobic fear reduction at immediate post-treatment assessment as well as at 9-month follow-up. The fact that EA PRS was associated with objective, real-world indicators of treatment success in this cohort, i.e. the number of dental appointments attended in
Figure 2. Top. Estimated mean polygenic scores of patients who were remitted at follow-up and those who were not (controlling for age, sex, population stratification and treatment type). Centre. Estimated mean polygenic scores of patients who were remitted at follow-up and those who were not in the multi-session treatment (controlling for age, sex and population stratification). Bottom. Estimated mean polygenic scores of patients who were remitted at follow-up and those who were not in the five-session treatment (controlling for age, sex and population stratification). EA: Educational Attainment; IQ: Intelligence quotient; NEUR: Neuroticism; SCHIZ: Schizophrenia; GAD: General Anxiety Disorder; MDD: Major Depression Disorder; BMI: Body Mass Index; error bars indicate standard errors of mean, * contains also patients who were not remitted at post-treatment assessment and did not attend follow-up assessment.
the post-treatment to follow-up interval, further underscores the predictive role of EA PRS in the context of this specific treatment. Moreover, in line with expectations we found lower GAD PRS in post-treatment remitters compared to non-remitters, as well as a negative influence of neuroticism PRS on post-treatment fear reduction, which suggests that genetic dispositions involved in the aetiology of anxiety disorders may also play an unfavourable role in treatment outcome. However, the influence of EA PRS identified in the dental fear cohort was much larger.

In relation to EA, our results support the notion that traits beyond cognitive ability alone might be associated with treatment outcomes, in that only the EA PRS – and not IQ PRS – was related to outcome in the dental fear cohort. EA PRS have previously been associated with both cognitive ability, as well as a broader suit of beneficial psychological traits, including self-control and interpersonal skills (Belsky et al. 2016). The contribution of traits beyond intelligence, such as self-efficacy, has also been demonstrated to account for the large heritability of educational attainment in a twin study (Krapohl et al. 2014). A very recent GWAS finding even showed that non-cognitive genetic factors, i.e. genetic variation in educational outcomes not explained by genetic variation in cognitive ability accounted for more than half of the genetic variance in EA and that heritable non-cognitive skills influence personality characteristics and downstream health outcomes (Demange et al. 2021). For example, genetics for non-cognitive skills were related to risk tolerance, delay of gratification as well as ‘Openness to Experience’, ‘Conscientiousness’ ‘Extraversion’ and ‘Agreeableness’ and negatively related to ‘Neuroticism’ (see Demange et al. 2021 for an overview). It is quite conceivable that participants with such a trait pattern may be more open to treatment offers and bravely and consistently implement newly learned behaviour, as, for example, we have observed in the dental treatment cohort in dealing with upcoming dental visits. Consequently, they may better benefit from treatment.

In contrast to our findings in the dental fear cohort, in our cohort of patients with mixed phobic fears, who were treated for about 20 exposure-based CBT sessions, we did not observe an influence of EA-PRS on treatment outcomes or any other gene effects. Ultimately, we can only speculate about possible reasons for this difference, but differences in cohort composition and treatment length may be possible reasons. Most patients in the dental treatment cohort initially sought dental surgery, and if critical dental fear levels were observed they were then referred for psychological treatment. Conversely, members of the mixed fear cohort had actively sought psychological help in a specialised treatment centre due to their mental health problems. It therefore seems plausible that the participants in these different cohorts might have different motivational starting points. This in turn could place higher demands during treatment on traits related to educational attainment in the dental fear cohort, such as disciplinary action control. Moreover, the mixed fear cohort had a higher education level on average and showed a tendency towards higher EA PRS compared to the dental fear cohort. The differential findings in relation to pre-treatment distress perhaps indicate a greater importance for EA PRS in the dental fear cohort. Further, the importance of dispositional prerequisites for EA in the dental fear cohort greatly increased in inverse relation to education levels, with EA PRS explaining 10% of variance between remitters and non-remitters amongst individuals without A-level equivalents, and less than 1% in more highly educated individuals. Due to these within- and between-cohort findings one may assume that genetic dispositions for EA are especially relevant in relatively less educated participants, both in relation to perceived mental distress and to treatment outcomes, and become less important amongst highly-educated individuals (although these findings could also reflect a lack of variance in the EA PRS scores amongst the more highly educated participants). Finally, there was a huge difference between the two cohorts in the length of treatment: In the dental fear cohort exposure was delivered for about three sessions, whereas at least 15 exposure sessions were administered in the mixed fear cohort. This could indicate that the importance of genetic dispositions, regardless of EA-related or not, can be ‘overwritten’ when treatment is of sufficient length. Of course, we cannot rule out that the difference in findings between the cohorts was due to the different disorders included (i.e. dental phobias vs other phobias) with different underlying psychopathology or differences in comorbidities between the cohorts. However, given the similarities of phobic disorders in relation to their phenomenology and resulting impairment (Muris and Merckelbach 2012) this possibility seems unlikely. The greater heterogeneity in the mixed phobic cohort (in terms of disorders) may have also led to greater heterogeneity in treatment outcomes and mechanisms, reducing the chance of finding a significant prediction by PRS scores. However, if so, this underscores the advantage of using relatively homogeneous
patient and treatment samples when aiming to investigate genetic influences on therapy outcomes.

Some limitations to our study need to be mentioned. Although we observed significant correlations between PRS and relevant trait indicators (DASS-21-scores) for most psychopathology-related PRS in our sample, the proportions of phenotypic variation explained by such polygenic traits in the original GWA was still low. When PRS is not well-powered to predict the target phenotype, correlations with the PRS can be misinterpreted and the results are vulnerable to type II errors. The absence of associations between PRS and other traits might be due to the sample size still being quite small for genetic analyses, at least for within-cohort analyses. Moreover, apart from the EA-related results, none of the other reported PRS outcome associations remain significant if adjustment for multiple testing is strictly applied in the regression analyses therefore requiring replication in an independent sample. Rather than relying on PRS derived from GWAS of related traits, a sufficiently powered GWAS of treatment outcomes holds some promise to provide a PRS that is more likely to have genetic overlap with treatment outcomes in our sample. Given that treatment outcomes are likely to be affected by many genetically influenced traits, it is perhaps less likely that any single trait polygenic score would capture enough genetic variance to be of clinical utility. As such, combining clinical samples with genetic and treatment outcome data in meta-analyses will be the first step towards developing a polygenic score for treatment outcomes. The largest GWA meta-analysis of psychological therapy was recently performed (\(N=2724\); Rayner et al. 2019), but no single genetic variants were detected, nor were there any polygenic score associations. It is likely that much larger samples will be required for both discovery and polygenic prediction.

To summarise, the reported findings suggest that individuals with relatively poor genetic propensity for EA-related traits may be a high risk group for poor psychological treatment outcomes, especially when combined with a relatively low level of education, and a psychological treatment that is short and ‘prescribed’ rather than sought. Future studies should therefore focus on how treatment should be tailored for the identified risk group in order to guarantee a therapy result that depends as little as possible on unfavourable genetic dispositions.

Acknowledgements

Christopher Rayner is funded by a studentship grant awarded to Thalia Eley and Gerome Breen by Foundation Peters. Thalia Eley is part-funded by a program grant from the UK Medical Research Council (MR/M021475/1). This study presents independent research [part-] funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. This research was funded by an Alexander von Humboldt professorship, awarded to Jürgen Margraf.

Disclosure of interest

The authors report no conflict of interest.

ORCID

André Wannemüller http://orcid.org/0000-0002-2602-986X
Robert Kumsta http://orcid.org/0000-0001-6006-6958
Thalia C. Eley http://orcid.org/0000-0001-6458-0700
Christopher Rayner http://orcid.org/0000-0003-2950-343X
Gerome Breen http://orcid.org/0000-0003-2053-1792
Jonathan Coleman http://orcid.org/0000-0002-6759-0944
Simon E. Blackwell http://orcid.org/0000-0002-3313-7084
Jürgen Margraf http://orcid.org/0000-0001-5207-7016

Data availability statement

The data that support the findings of this study are available from the corresponding author, AW, upon reasonable request.

References


Schneider S, Margraf J. 2006. Diagnostisches Interview bei psychischen Störungen (DIPS für DSM-IV-TR) [Diagnostic Interview for Mental Disorders]. Berlin: Springer.


